DISSECTING BCR-ABL VARIANT SIGNALING PATHWAYS USING NOVEL INTERACTOME IDENTIFICATION STRATEGIES

By

Jevon A. Cutler

A dissertation submitted to The Johns Hopkins University in conformity with the requirement of the degree of Doctor of Philosophy

Baltimore, MD

May 2018

ABSTRACT

Cell signaling is an essential function of cells and tissues. Understanding cell signaling necessitates technologies that can identify protein-protein interactions as well as post translational modifications to proteins within protein complexes. **The goals of this study are (1) to understand how BCR-ABL variants differentially signal to produce different clinical/experimental phenotypes and (2) to develop novel interactome detection strategies to understand signaling.** This dissertation describes an integrated approach of the use of proximity dependent labeling protein-protein interaction analysis assays coupled with global phosphorylation analysis to investigate the differences in signaling between two variants the oncogenic fusion protein, BCR-ABL. Two major types of leukemogenic BCR-ABL fusion proteins are $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$. Although the two fusion proteins are closely related, they can lead to different clinical outcomes. A thorough understanding of the signaling programs employed by these two fusion proteins is necessary to explain these clinical differences. Our findings suggest that $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ differentially activate important signaling pathways, such as JAK-STAT, and engage with molecules that indicate interaction with different subcellular compartments. In the case of $p210^{BCR\text{-}ABL}$, we observed an increased engagement of molecules active proximal to the membrane and in the case of $p190^{BCR\text{-}ABL}$, an engagement of molecules of the cytoskeleton. These differences in signaling could underlie the distinct leukemogenic process induced by these two protein variants. Additionally, this dissertation also describes the development of a novel interactome detection strategy, called Biotinylation Site Identification Technology (BioSITe), which increases the sensitivity and specificity of proximity dependent biotin labeling technologies. When applied to BCR-ABL variants,

BioSITe provides structural information about BCR-ABL interacting proteins and the degree of proximity these proteins are to BCR-ABL. Finally, this thesis demonstrates the use of isotopically labeled biotin for quantitative BioSITe experiments, applied to BCR-ABL variants, simplifies differential interactome analysis.

DISSERTATION REFEREES

Graduate Advisor: **Karen Reddy Ph.D..,** Professor, Department of Biological Chemistry and Center for Epigenetics

Graduate Advisor: **Akhilesh Pandey, M.D. Ph.D.,** Professor, McKusick-Nathans Institute of Genetic Medicine, Departments of Biological Chemistry, Oncology, and Pathology

ACKNOWLEDGEMENTS

This dissertation is the product of seven years of intensive research and molecular biology and genetics education which exemplifies the multidisciplinary nature of the Human Genetics training program at the Johns Hopkins University School of Medicine. The Human Genetics training program curriculum has provided me a strong foundation in one of the few true fundamentals of biology, genetics. Building upon the unique pedagogy cultivated by the faculty within the Institute of Genetic medicine whose goal is to understand human genetic variation and disease I have become a proud and skillful scientist. Of course, my two incredible thesis mentors, Karen Reddy and Akhilesh Pandey, have had the most intimate hand in shaping my experience of obtaining my PhD by providing an immense amount of material support, nourishment as a thinking scientist, and technical training. Certainly, co-mentorship has its challenges and while the interests of the Reddy and Pandey labs are disparate in the end Karen and Akhilesh put me in a privileged position to influence the direction of both labs. Both Karen and Akhilesh gave me wide room to explore many different aspects of molecular biology and stuck with me when things didn't work and helped me flourish when they did. Being in two labs has given me the opportunity to learn a more diverse set of technologies that I wouldn't have overwise in only one lab. Most of all both Karen and Akhilesh have taught me how to critically think about and pursue scientific questions. Both approach science in different ways and it was an immensely great experience to benefit from these different approaches. I truly admire both Karen and Akhilesh for their innovative and forward thinking scientific pursuits and their openness toward collaborations in science, for which

I have benefited immensely from. I look forward to continue that collaborative spirit throughout my career.

A special thanks to Dr. Oliver Hantschel and Dr. Reckel Sina Maren who were working on the same BCR-ABL signaling story as myself and had the humility to coordinate our efforts so that we could publish back-to-back in Leukemia. It was an incredible experience to work with Sina and Oliver and they ended up co-authors on the BioSITe paper, contributing an important analysis. Another special thanks to Dr. Dae In Kim who I worked very closely with to produce the BioSITe paper. It was a wonderful collaborative experience to work with Dae In. I'd also like to thank Dr. Curt Civin who has been a mentor from afar (across town at University of Maryland) from the first month I arrived in Baltimore Curt has been an invaluable resource for navigating science, Baltimore, and beyond. Curt was instrumental in helping me find a post doctorate position and has heaped so much praise upon me that has lifted my spirits and sense of self on many occasions. I'd also like to thank Curt for initiating the collaboration with a colleague of his Dr. Tami Kingsbury, who I've gotten the opportunity to work on her exciting project.

I would like to thank my thesis committee, Dr. Steven Desiderio, Dr. Hans Bjornsson, and Dr. Patrick Brown, for all of their valuable insight and guidance over the course of my PhD. I am especially grateful to the generosity, support, and friendship I received from previous and current lab members of the Reddy and Pandey lab. Members including Tai-Chung Haung, Min-Sik Kim, Chris Mitchell, Teresa Luperchio, Xianrong Wong, Mo Heydarian, Raiha Tahir, Xinyan Wu, Saddiq Zahari, Tori Hoskins, Chan-Hyun Na, and many others. I would like to especially thank my undergraduate research

mentors, Dr. Michael Loken and Dr. Denise Wells, for recognizing my potential to become a scientist and fomenting my interest in hematopoiesis.

I would also like to thank my incredibly supportive family of Denise, Mike, Harvey, Roselyn, Krista, Sara, and Zach. I'd also like to thank Rachel, the mother of our two children Veja and Ivo, for being game to move to Baltimore and start a family with me, during the trying times of graduate school. Rachel's day-to-day presence was probably the single greatest impact on my happiness. Of course, Veja and Ivo also share in that as they are the most sweet and wonderful little bubs ever.

Many thanks to Sandy Muscelli for her excellent skills as an administrator, her kindness, sense of humor, and the many pretzels she provided me. Sandy's support throughout the years has been invaluable and her allowance of a very reasonable part of the program budget for me to organize some great happy hours has been wonderful. I'd like to thank Dr. Dave Valle for his incredible leadership of the Human Genetics Training Program, his excellence in teaching genetic medicine and his valuable historical knowledge of the field of genetics. I'd also like to thank the Department of Biological Chemistry for their support during my graduate career, especially Dr. Gerry Hart's commitment to funding my salary.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 3

Chapter 4

Chapter 5

LIST OF ABBREVIATIONS

ABL	Abelson tyrosine kinase
Abi1	Abl interactor 1
Acap1	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 1
Arap1	Arf-GAP with a rho-GAP domain, ANK repeat and PH domain-containing protein 1
Arap3	Arf-GAP with a rho-GAP domain, ANK repeat and PH domain-containing protein 3
Arhgef6	rho guanine nucleotide exchange factor 6
Arhgdia	rho GDP-dissociation inhibitor 1
Arid3a	AT-rich interactive domain-containing protein 3A
BCR	breakpoint cluster region
BioSITe	Biotin Site Identification Technology
B-ALL	B cell acute lymphoblastic leukemia
Cbl	Cbl proto-oncogene
Cblb	Casitas B-lineage lymphoma b
Clnk	cytokine dependent hematopoietic cell linker
CML	chronic myelogenous leukemia
Crk	CRK proto-oncogene adaptor protein
Csk	tyrosine protein kinase Csk
Dab2ip	DAB2 interacting protein
DH	Dbl homology

DiDBiT	Direct Detection of Biotin-containing Tags
Dok1	Docking protein 1
E2A	transcription factor E2-alpha
GAPs	GTPase-activating proteins
Gab2	Grb2-associated-binding protein 2
GEF	guanine nucleotide exchange factors
GFP	green fluorescent protein
Grb2	growth factor receptor-bound protein 2
IRES	internal ribosome entry site
HSC	hematopoietic stem cell
LC-MS/MS	liquid chromatography and tandem mass spectrometry
Lyn	tyrosine protein kinase Lyn
Lrrk1	leucine rich repeat kinase 1
MLL-r	mixed lineage leukemia rearrangements
Nck1	Nck adaptor protein 1
Nck2	Nck adaptor protein 2
Peak1	pseudopodium-enriched atypical kinase 1
PH	Pleckstrin Homology domain
PTM	post translational modification
PPI	protein-protein interaction
PDLTs	proximity dependent labeling technologies
pTyr	phosphotyrosine
Ptpn11	Tyrosine-protein phosphatase non-receptor type 11

RUNX1	runt related transcription factor 1
RUNX3	runt related transcription factor 3
Sch1	SHC-transforming protein 1
shRNA	short hairpin RNA
SHIP1	phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1
SHIP2	inositol Polyphosphate Phosphatase Like 1
Sipa1	signal-induced proliferation-associated protein 1
Sia1l11	signal-induced proliferation-associated 1-like protein 1
STAT	signal transducer and activator of transcription
Sos1	son of sevenless homolog 1
Sos2	son of sevenless homolog 2
Syk	tyrosine protein kinase Syk
Tnk2	tyrosine kinase non-receptor 2
Ubash3b	ubiquitin-associated and SH3 domain-containing protein B
Vav1	proto-oncogene Vav 1
WAS	Wiskott aldrich syndrome
WASP	Wiskott aldrich syndrome protein
Wasf2	WAS Protein Family Member 2
Wipf1	WAS/WASL-interacting protein family member 1
Wipf2	WAS/WASL-interacting protein family member 1
Zyx	focal adhesion localized protein Zyxin

CHAPTER 1.

INTRODUCTION

Cell signaling is an essential function of cells and tissues. Cells respond to environmental signals and via signal transduction translate these signal into the augmentation of cellular processes to carry out important biological functions. The molecular events that underline signal transduction interest scientists and clinicians because many times these events are abnormally altered in the state of disease. Furthermore, these molecular events hold promise to be pharmacologically targeted to provide important therapies. The study of cell signaling has a long history of innovation and breakthroughs[1]. Undoubtedly, one of the most studied molecular events that mediate signal transduction has been the regulation of post translation modifications (PTMs) of proteins. While there are close to 200 different known PTMs, phosphorylation is perhaps the most studied and importantly appears to be involved in all cellular processes. Phosphorylation is a reversible modification to the serine, threonine and tyrosine amino acids on proteins that is mediated by kinases and phosphatases that phosphorylate and dephosphorylate, respectively, target proteins. The cascade of phosphorylation/dephosphorylation events allow cells to transmit signals from environmental stimulus to alter the function of proteins that control important cellular functions that range from cytoskeleton reorganization to changes in gene expression (Figure 1.1).

Another essential molecular event that governs signal transduction if protein-protein interactions (PPIs). The function of an individual protein cannot be understood without an understanding of what other proteins it is in close proximity to or physically interacts with. For example, a kinase phosphorylates substrates and this is mediated by not only

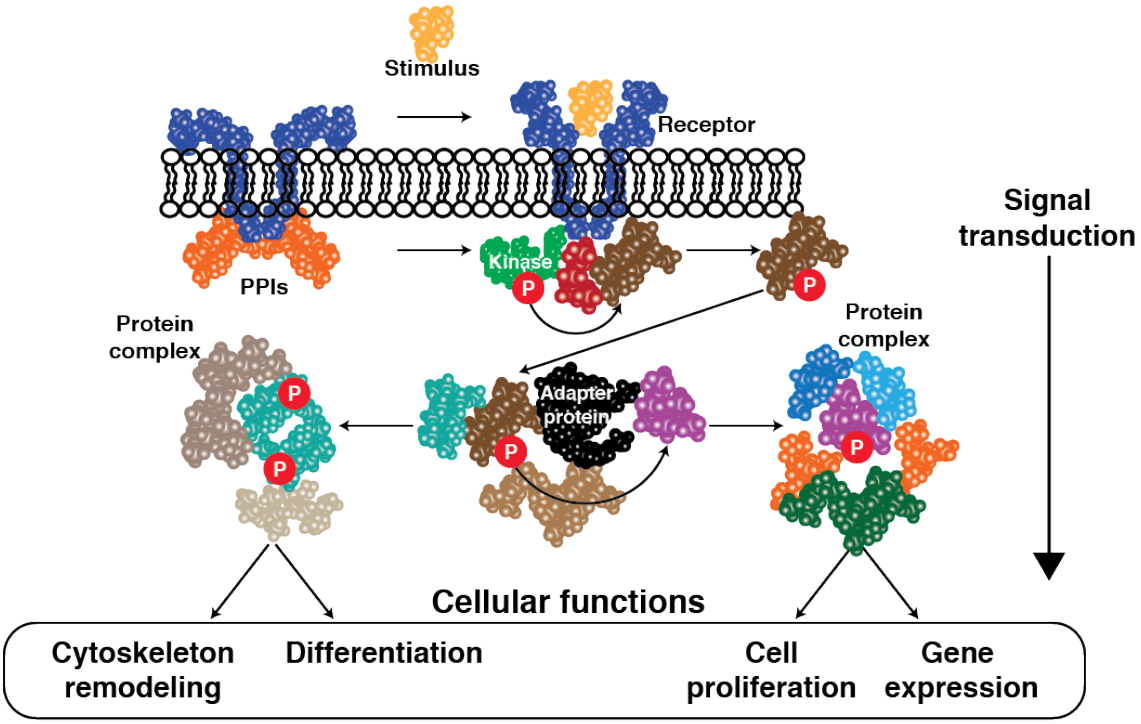

Figure 1.1 Simplified model of signal transduction mediated by PPIs. Depiction signal transduction in response to a stimulus, in this case a ligand binding to a receptor. A generalized receptor shown in dark blue binds ligand (yellow) causes a conformational change allowing new PPIs that bring together kinases, green and brown, and substrates, red and brown. Subsequently, the activated kinases produce a phosphorylation, red circle labeled P, cascade that transmits downstream phosphorylation events that affect protein functions that result in altered cellular functions. Protein complexes mediating signal transduction are depicted by closely place proteins.

an affinity for the protein it can phosphorylate but also the proteins that aid increasing the efficiency of this reaction, often referred to as adapter proteins. Recently, advances in the study of PPIs using in-cell proximity dependent labeling technologies (PDLTs) have shown promise to change the conventional understanding of PPIs. Based on the simple concept of tethering a engineered biotin ligase to a protein of interest and expressing this fusion protein, PDLTs such as the BioID and APEX systems are being rapidly deployed by researchers to study PPIs as well as subcellular proteome mapping[2]. Both the BioID and APEX systems leverage biotin ligases that produce short lived and highly reactive biotin adducts that only diffuse a short distance from the active site of the ligase giving both systems the ability to selectively label only proximal proteins. The APEX system leverages an engineered ascorbate peroxidase enzyme designated APEX2 that uses hydrogen peroxide to catalyze the transfer of biotin-phenol to tyrosine residues in proteins[3,4]. BioID, similarly relies on an engineered biotin ligase called BirA*[5,6] to generate biotinoyl-5′-AMP which is reactive with lysine residues of proteins [7]. The labeling time is starkly different in the BioID systems compared to the APEX system, BioID requires 18-24 hours of exogenous biotin incubation to achieve complete labeling and the APEX system can modify proximal proteins with exogenously provided biotin phenol in 30 seconds to 1 minute of labeling time (Figure 1.2). These systems have three important advantages to traditional methods (1) labeling is controllable by adding exogenous biotin (2) labeling is carried out under cellular conditions with sub-cellular structures intact providing more biologically relevant conditions with no in-vitro bias (3) sample preparation can be made uniform and reproducibly carried out lab-to-lab. When compared to traditional methods to study PPIs it is clear that proximity dependent

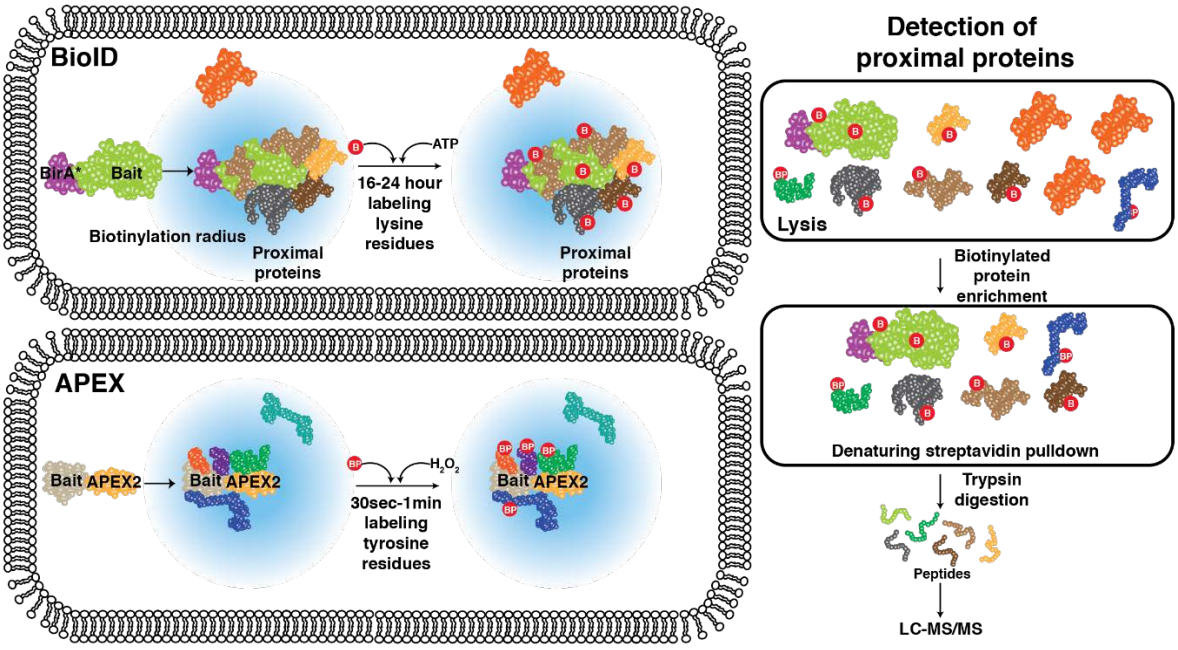

Figure 1.2 Overview of proximity dependent labeling technologies. Depiction the BioID and APEX systems to detect proximal proteins using in cell biotin labeling technologies. BioID system, top left, relies on the engineered biotin ligase BirA*, which requires exogenous biotin (B in red circle) and ATP to biotinylate proximal proteins in and 18-24 hour period. APEX system, bottom left, relies on the engineered biotin ligase APEX2, which requires exogenous biotin phenol (BP in red circle) and hydrogen peroxide (H_2O_2) to biotinylate proximal proteins. A generalized workflow for identifying proximal proteins in both APEX and BioID systems is depicted on the right. Cells harboring APEX and BioID systems are lysed and enriched for biotinylated proteins and subsequently digested into peptides for LC-MS/MS analysis. Biotinylated proteins are labelled with B or BP in red circles and non-biotinylated proteins are removed with strong denaturing conditions.

labelling technologies detect a distinct, “complimentary” and “biologically meaningful” subset of the interactions of the protein of interest[8].

Both the APEX and BioID system have been coupled to liquid chromatography and tandem mass spectrometry (LC-MS/MS) as the mode of detection of interacting proteins. Generally speaking, the study of protein-protein interactions and phosphorylation has taken a major leap forward with the advent of so-called “shotgun proteomics” using LC-MS/MS[9]. Where historically studying protein involved low throughput methods, mostly antibody based, to detect single proteins and single PTMs, LC-MS/MS allows for high-throughput simultaneous detection of thousands of proteins and PTMs. Shotgun proteomic approaches rely on enzymatic, usually trypsin, digestion of proteins into peptides that can be identified by LC-MS/MS followed by database search of resulting mass spectra.

The broad biological interest of my thesis work is rooted in the hematopoietic system and the leukemias that arise from this system. More specifically, my thesis work focuses on how signaling drives the leukemic process. The hematopoietic system, simply put, is the responsible for the formation of cells that comprise blood. The hematopoietic stem cell (HSC) resides in the bone marrow and gives rise to many different cell types that are responsible for incredibly diverse functions throughout the body. The hematopoietic system is constantly active, meaning cell divisions is a constant producing hundreds of billions cells a day. A subset of these cells that are on my interest to my thesis work are lymphocytes. Lymphocytes in the presence of genetic mutations can produce a wide variety of cancers including, and of most interest to my thesis work, leukemia. Leukemia can arise from many different stages of maturation from the HSC and can retain

phenotypes characteristics of normal cells giving rise to nomenclature that helps clinicians stratify and treat these various types of leukemia (Figure 1.3). One such subset of leukemia, B cell acute lymphoblastic leukemia (B-ALL), is of particular interest to my thesis work because of many of the underlining genetic mutations produce fusion proteins that have fascinating biology and important clinician characteristics that pose challenges to treat. My thesis work focuses on one such mutation that results from a chromosomal translocation that produces the fusion gene known as BCR-ABL. BCR-ABL is most famously associated chronic myelogenous leukemia (CML), which has very good outcome because of targeted treatments (cite), but is also associated with B-ALL which has inferior prognosis.

Throughout hematopoietic development differentiating cells respond to cytokines and other signals that provide the stimulus for correct development. These signaling pathways are subverted and/or altered in leukemic cells which underlines a major mechanism of leukemogenesis and leukemic maintenance[10]. In *Chapter 2* (methods) and *Chapter 3* I will describe a set of experiments using the BioID system and global phosphorylation profiling to interrogate outstanding questions related how the BCR-ABL oncogene through signaling can differently generate B-ALL and CML. Chromosomal translocations involving the breakpoint cluster region (BCR) and Abelson murine leukemia viral oncogene homolog 1 (ABL) genes generate BCR-ABL, a potent oncogene that drives leukemia. Depending upon the exact breakpoint of translocation within the BCR gene, distinct fusion genes leading to different fusion proteins are generated. The most common translocations generate fusion protein products with molecular weights of 190 and 210 kD.[11]

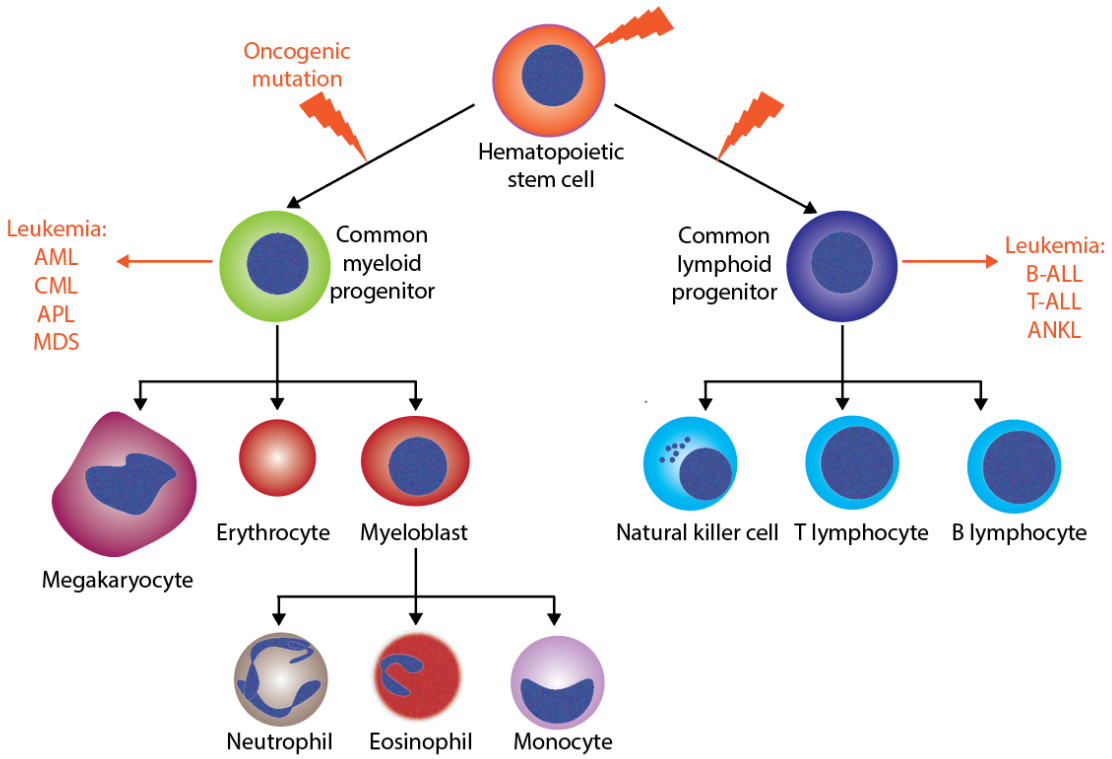

Figure 1.3 Simplified schematic of hematopoiesis and leukemogenesis. Depiction of different stages of hematopoietic maturation derived from the hematopoietic stem cell (HSC). Orange lightning strikes represent DNA mutations to either progenitor or HSC populations that give rise to leukemia. Examples of leukemia with phenotypic characteristics of different lineages are listed, acute myeloid leukemia (AML), chronic myelogenous leukemia (CML), Acute promyelocytic leukemia (APL), myelodysplastic syndromes, B-cell acute lymphoblastic leukemia (B-ALL), T-cell acute lymphoblastic leukemia (T-ALL), and aggressive NK-cell leukemia (ANKL).

These variants, designated as p190 and p210 (hereafter referred to as $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$), are associated with distinct clinical features.[12] Generally, $p190^{BCR\text{-}ABL}$ leads to ALL while the $p210^{BCR\text{-}ABL}$ variant leads to CML. The two types of leukemia driven by these variants exhibit different treatment response to chemotherapy including tyrosine kinase inhibitors leading to distinct clinical outcomes.[13,14] The reasons for this discrepancy are still not fully understood.

BCR-ABL fusion genes code for proteins with constitutively active tyrosine kinase activity that mediate transformation by inducing signaling pathways that enhance cell growth and evade normal regulation. Despite numerous studies into the mechanisms underlying $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ induced leukemia, the fundamental questions of whether and how these proteins engage different signaling programs remain unanswered. The major structural difference between the two variants is the presence two domains commonly found in guanine nucleotide exchange factors (GEFs); a Dbl homology (DH) domain followed by a Pleckstrin homology domain (PH) in $p210^{BCR\text{-}ABL}$. Dbl homology domains have catalytic GEF activity and PH domains facilitate intercellular targeting including plasma membrane binding. Ablation of the DH domain's GEF activity or deletion of the PH domain results in phenotypes similar to $p190^{BCR\text{-}ABL}$, suggesting that both domains may be essential for sustaining $p210^{BCR\text{-}ABL}$-specific phenotypes.[12,15] Direct comparison of $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ in *in vitro* kinase activity assays has shown $p190^{BCR\text{-}ABL}$ to exhibit higher auto-phosphorylation, leading to the conclusion that $p190^{BCR\text{-}ABL}$ is likely a more active kinase.[16,17] Several studies have analyzed signaling pathways induced by BCR-ABL (reviewed in Hantschel[18] and Cilloni et al.[19]) although, studies directly comparing signaling differences between $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ are more

limited.[12,20,21] A few studies have examined global phosphorylation events induced by either p190$^{BCR\text{-}ABL}$ or p210$^{BCR\text{-}ABL}$,[22–24] but to date no study has directly compared the differential signaling patterns induced by these two variants in the same cellular environment. Similarly, very little is known about the potential protein-protein interaction differences that mediate p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ specific signaling. The p210$^{BCR\text{-}ABL}$ "core" interactome has been previously catalogued[25] but, to date, there are no reports of direct comparisons between p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ interactomes.

In order to test the hypothesis that p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ leverage different protein partners and induce different signaling programs, we decided to express these proteins in the same cell type. We chose an early hematopoietic progenitor model, the well-studied Ba/F3 cell line, to directly compare the differential signaling and protein-protein interactome of the two variants. We used the BioID system to identify protein-protein interactions and phosphotyrosine (pTyr) peptide enrichment to identify site specific tyrosine phosphorylation events on target proteins. An integrative analysis of the interactome and phosphorylation data allowed us to discover overlapping and differential signaling pathways. Specifically, we found an increased interaction of p210$^{BCR\text{-}ABL}$ with membrane-proximal signaling proteins including ubiquitin-associated and SH3 domain-containing protein B (Ubash3b or Sts-1) and phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1 (Inpp5d or SHIP1), with both proteins displaying increased phosphorylation in p210$^{BCR\text{-}ABL}$ harboring cells. p190$^{BCR\text{-}ABL}$, on the other hand, displayed increased interaction and phosphorylation of many cytoskeletal proteins, including members of the Wiskott-Aldrich syndrome protein family. In agreement with previous studies, we also observed differential p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ phosphorylation signatures of signal

transducer and activator of transcription (Stat) proteins. In addition, our study provides evidence for differential phosphorylation of Stat proteins at the level of individual tyrosine phosphorylation sites, which could be mediated by differential interactions. Our findings reveal that $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ have both overlapping and different specificities for substrates, potentially interact with different subcellular compartments and mediate differential downstream signaling events that ultimately lead to different clinical outcomes.

As described above PPIs are essential to understand the function of proteins. Developing novel methods to detect PPIs is of great importance. Building upon the work of integrated analysis of BCR-ABL variant interactome and phosphorylation profiles, *Chapter 2* (methods) and *Chapter 4* describe the development of a method to detect biotinylated peptides to improve upon PDLTs. Mapping PPIs and subcellular proteomes is an important strategy to reveal biological phenomena and mechanisms of disease. As described above the newly developed proximity-dependent biotinylation methods such as APEX and BioID facilitate discovery of PPIs and subcellular proteomes within living cells. The identification of biotinylated proteins that are generated using these strategies has hitherto relied upon capture using the streptavidin family of proteins followed by either on-bead digestion or in-gel digestion followed by LC-MS/MS analysis. While these methods have been used successfully they both necessitate the use of identified non-biotinylated peptides as surrogates for biotinylation, as biotinylated peptides are rarely, if at all, detected. The reason non-biotinylated peptides are rarely identified is that the way that the strong affinity of streptavidin for biotin is used largely precludes it. In in-gel digestion elution of biotinylated proteins is required. Several elution methods have been developed, either

using detergent[26], extremely low pH[27] or solvents[28] and alternative strategies have been employed including weakened affinity of streptavidin/avidin to biotin through chemical[29] or genetic[30] means have also been described. However, all approaches yield partial recovery of biotinylated proteins from streptavidin[31] and when elution followed by in-gel digestion is applied in proximity-dependent biotinylation experiments non-biotinylated peptides vastly out number biotinylated peptides[3]. The need for good elution strategies has been circumvented by the use of on-bead digestion which is a convenient and widely used method in proximity-dependent biotinylation experiments. However, biotinylated peptides are left bound to beads and remain undetected as the streptavidin-biotin interaction is not disrupted. In a different approach, which does not suffer from the limitations just described, Schiapparelli *et al.* previously described a strategy designated DiDBiT (Direct Detection of Biotin-containing Tags) where they first digested proteins into peptides and then enriched biotinylated peptides using NeutrAvidin[32]. Because this method was mostly tested using proteins biotinylated *in vitro* we reasoned that the strong affinity of NeutrAvidin for biotin could still limit detection of biotinylated peptides, especially in the context of *in vivo* biotinylation such as in proximity-dependent biotinylation methods, where the overall level of biotinylation level might be low.

To improve the biotinylated peptide detection issue, we have developed a strategy designated **Bio**tinylation **S**ite **I**dentification **Te**chnology (**BioSITe**), which is based on the use of anti-biotin antibodies to directly capture and identify biotinylated peptides in a single LC-MS/MS run. We developed BioSITe using BCR-ABL variant system described in *Chapter 3* but also apply it to a well-established APEX based system of subcellular proteome mapping of compartments of the mitochondria. We show that

detection of biotinylated peptides greatly increases the confidence of candidate protein identification in proximity-dependent biotinylation methods. By providing the site of biotinylation on proteins labeled in these methods, this approach also offers a new level of information about the structural aspects of PPIs and protein topology. Finally, we describe a simple approach for quantitative BioSITe experiments through the use of isotopically labeled biotin obviating the use of metabolic labeling based quantitative strategies such as SILAC[33].

CHAPTER 2.

MATERIALS AND METHODS

2.1 Plasmids, cloning, antibodies, and reagents

The mouse retroviral vector pMSCV modified with an IRES-GFP containing either the $p190^{BCR\text{-}ABL}$ or the $p210^{BCR\text{-}ABL}$ form of BCR-ABL was provided by Saradhi Mallampati and Xiaoping Sun. The plasmid pcDNA3.1 myc-BioID containing the mutant BirA-R118G was purchased through Addgene (Plasmid #35700). BirA was cloned in-frame on the N-terminus of both $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ to generate BirA-R118G-$p190^{BCR\text{-}ABL}$-IRES-GFP-pMSCV and BirA-R118G-$p210^{BCR\text{-}ABL}$-IRES-GFP-pMSCV. A fragment between the AUG start codon and an XhoI site within the common part of BCR in $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ was PCR amplified with flanking restriction sites for 5' XhoI-HpaI and 3' SalI-XhoI. Parent plasmids were restriction digested with to XhoI to remove 147 bases upstream of the AUG, the PCR product was ligated in its place. The $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ plasmids now missing 147 base pairs 5' of the AUG were digested with blunt end restriction enzyme HpaI. The plasmid pcDNA3.1 myc-BioID containing the mutant BirA-R118G was purchased through Addgene (Plasmid #35700). The plasmid pcDNA3.1 myc-BioID containing the mutant BirA-R118G was purchased through Addgene (Plasmid #35700) was restriction digested with NheI and BamHI to excise BirA-R118G. This fragment was blunt end cloned into the HpaI site of $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ making, in-frame, BirA-R118G- $p190^{BCR\text{-}ABL}$-IRES-GFP-pMSCV and BirA-R118G-$p210^{BCR\text{-}ABL}$-IRES-GFP-pMSCV. SILAC amino acids were purchased from Cambridge Isotopes; heavy (lysine-$^{13}C_6$,$^{15}N_2$ #CNLM-291, arginine-$^{13}C_6$,$^{15}N_4$ #CNLM-539) and medium (lysine-$^{13}C_4$ #CLM-2247, arginine-$^{13}C_6$ #CLM-2265) labels were dissolved in arginine and lysine deficient RPMI 1640 medium, then sterile filtered. A complete list of antibodies used: α-ABL1 (emdMillipore #MABT203), α-WIP/WIPF1

(emdMillipore #MABS1275), α-WASP (Cell Signaling Technology(CST) #4860), α-phospho-WASP (Tyr290) (Abcam #59278), α-Ubash3b/STS-1 (Santa Cruz Biotech. sc-514612), α-SHIP1 (CST #2728), α-Stat1 (CST #9172), α-phospho-Stat1 (Tyr701) (CST #7649), α-Stat5b (Santa Cruz Biotech. sc-1656), α-phospho-Stat5 (Tyr694) (CST #9359), α-Stat6 (Santa Cruz Biotech. sc-374021), α-phospho-STAT6 (Tyr641) (Abcam #125308), α-DOK1 (Santa Cruz Biotech. sc-6929), α-phospho-DOK1 (Tyr362) (Abcam #47374), α-Cbl (Santa Cruz Biotech. sc-170), and α-β-Actin (CST #4970) Anti-biotin antibody 1 (Abcam, #ab53494), anti-biotin antibody 2 (Bethyl Laboratories,#150-109A), streptavidin-HRP (Abcam, #ab7403), protein G beads (EMD Millipore, #16-266), high capacity neutrAvidin agarose (Thermo Fisher Scientific, #29202), biotin (Sigma Aldrich, #B4501), biotin-2',2',3',3'-d_4 (Sigma Aldrich, #809068), lipofectamine 2000 (Thermo Fisher Scientific, #11668019), biotin-phenol (Iris Biotech, #CDX-B0270-M100), sequencing grade trypsin (Promega, #V5113), GalT1 enzymatic labeling kit (Invitrogen, #C33368), PNGase F (New England Biolabs, #P0704S), (Thermo Fisher Scientific, #C10412) and trypsin (Worthington Biochemical Corporation, #LS003741).

2.2 BioID: biotinylated protein capture, on bead digestion, and Western blotting validation:

Methods used for execution of the BioID experiment were an amalgamation of protocols from the Roux and Gingras groups[8,34] with slight modifications to both. Cells destined for BioID protocols were cultured for 24 hours in 50 μM biotin containing culture medium. Lysis was carried out in BioID lysis buffer (50 mM Tris, pH 7.5, 500 mM NaCl, 0.4% SDS, 2% Triton X-100 with Halt™ protease inhibitors) by probe sonication as described above. After the first round of sonication, the sample was diluted with equal volume of

50mM Tris, target protein concentration was 5 mg/ml. Lysates were quantified by BCA and dithiotheritol (DTT) was added to a final concentration of 1 mM for the pulldown. A ratio of 4 mg of lysate per 300uls of MyOne™ Dynabeads® Streptavidin C1 (#65001) to 5 µg of trypsin per replicate was used for all mass spectrometry experiments. For label-free experiments 4 mg per IP replicate was used. For SILAC experiments 8 mg per cell line was mixed. An aliquot of input was collected for western blot analysis and lysate/bead slurry was rotated overnight at 4°C, aliquot of the flow-through was collected. All bead washes were in 1.3 mL; 2 times with 2% SDS, 50 mM Tris, pH 7.5 followed by 2 washes with bioID lysis buffer, and finally 3 washes of 50 mM ammonium bicarbonate (ABC). On the third wash 10% of the total bead slurry was aliquot for boiling elution with 1X Laemmli Sample Buffer (Bio-Rad #1610747) with 20mM DTT for immunoblotting. Additional streptavidin isolates were generated by elution for validation of mass spectrometry data by immunoblotting. The rest of the bead slurry was subjected to on-bead digestion. Wash buffer was removed and beads were spun down and resuspended in 50 µl of 8 M UREA, 50 mM ABC, pH 8.0 to denature the bound proteins. Proteins were reduced by adding 5 µl of 1 M DTT and incubating at 30°C for 60 minutes, then samples were alkylation by adding 5 µl of 0.5M iodoacetamide in the dark for 30 minutes. The volume was then increased to 400 µl with 50 mM ABC pH 8.0 containing 5 µg of sequencing grade trypsin. Digestion was allowed to go overnight on a shaker kept at 30°C. The beads were removed from the digest solution and washed once with 50 µl of 50 mM ABC and all beads removed, the solution was pooled. The final 450 µl was acidified to 1% trifluoroacetic acid, spun at 18,000xg and loaded on to a C_{18} homemade stage tip for clean-up, stage-protocol was based on previously published methods.

2.3 Cell Culture and generation of stable cell lines:

Mouse retroviral packaging cell line Platinum-E was cultured in high-Glucose Dulbecco's modified Eagle's medium (Gibco #11995073) supplemented with 10% FBS (ThermoFisher #26140079). The mouse hematopoietic cell line Ba/F3 was cultured in RPMI 1640 (ThermoFisher #11875119) supplemented with 10% FBS in the presence of the cytokine IL-3 (0.5 ng/mL). The mouse MPP cell line generated from *Ebf1*-/- were cultured on OP-9 stromal cells in Opti-MEM I Reduced Serum Medium (ThermoFisher #31985088) supplemented with 5% FBS, 1% penicillin-streptomycin, 2 mM glutamine, beta/mercaptoethanol (1:1000 dilution). MPP cultures were grown in the presence of IL-7 (5 ng/mL; CST #5217SC), Scf (10 ng/mL; CST #5223SC), and Flt3l (10 ng/mL; R&D systems #427-FL-025)[35,36].

Plasmid DNA from either p190-IRES-GFP-pMSCV, p210-IRES-GFP-pMSCV, BirA-R118G-p190-IRES-GFP-pMSCV, or BirA-R118G-p210-IRES-GFP-pMSCV was transfected into Platinum-E using the Fugene HD Transfection Reagent (Promega #E2311). Viral supernatant was collected at multiple time points and pooled. Ba/F3 and MPP cell lines were subjugated to spinfection mediated retroviral transduction. Two million cells from each cell line were resuspended in 1 mL of viral supernatant with cytokines and polybrene (1 μg/mL) then centrifuged in a 24-well plate at 2000xg for 2 hours at 24°C. The cells were returned to a 37°C incubation for a 2-hour recovery. Cells were then washed with PBS and returned to normal cytokine containing medium and in the case of MPP cells were returned to OP-9 stromal cells. Following a 24 hour incubation transduction efficiency was confirmed using flow cytometry to quantify GFP expression. Cells were then washed with PBS to remove all cytokines and returned to culture

conditions without cytokines or stromal cells to select for growth factor independence. MPP cells expressing $p190^{BCR\text{-}ABL}$ or $p210^{BCR\text{-}ABL}$ were transitioned into SILAC RPMI 1640 (ThermoFisher #89984), for pTyr experiment, Ba/F3 BirA-$p190^{BCR\text{-}ABL}$, BirA-$p210^{BCR\text{-}ABL}$, and parental cells were also transitioned to SILAC RPMI, otherwise all cells were grown in their respective medium. All cell lines when used in proteomic analysis were washed in large volumes of PBS at least 3 times and pelleted and snap frozen with liquid nitrogen.

2.4 Peptide preparation and phosphopeptide enrichment:

Frozen SILAC Labeled cell pellets were lysed in 9 M urea buffer with protease inhibitors (Halt™ Cocktail ThermoFisher #78429), sodium pyrophosphate (2.5 mM), sodium orthovanadate (1 mM), sodium fluoride (5 mM), β-glycerophosphate (1 mM) by 3 rounds of probe sonication (duty cycle 30%, 20 two second pulses). Protein concentration of lysates was determined by Pierce™ BCA Protein Assay Kit. Twenty milligrams of protein lysates from each labeling condition was mixed and subjected to trypsin in-solution digestion, followed by lyophilization. A total of 60mg and 40 mg for the BAF3 and MPP experiment, respectively, per replicate of lyophilized tryptic peptides were used for phosphotyrosine enrichment using PTMScan® Phospho-Tyrosine rabbit monoclonal antibody kit (P-Tyr-1000) (Cell Signaling Technology, #8803). Anti-phosphotyrosine immunoprecipitation was carried out following the manufacturer's instructions. A 300 µg aliquot of SILAC mix was also subjected to a separate in-solution tryptic digestion and was fractionated by basic reverse phase liquid chromatography into 12 fractions and analyzed by LC-MS/MS. For BioSITe and other peptide enrichment experiments, cells were lysed using a lysis buffer containing 50 mM TEABC and 8 M Urea and sonication

as described above. The protein concentration of samples was measured by BCA assay and the protein concentration of APEX lysate was estimated by cell number. Reduction and alkylation were performed by serial incubation with 10 mM DTT for 30 min and by 20 mM IAA for 30 min in the dark. Lysate was diluted to 2 M Urea by adding three cell lysate volumes of 50 mM TEABC. The proteins were digested with trypsin (1:20 of trypsin to protein) at 37°C overnight. The resulting tryptic peptides for both experiments were desalted using a Sep-PAK C_{18} and subsequently lyophilized.

For on-bead digest, neutravidin beads were centrifuged and resuspended in 100 µL of 50 mM TEABC and 2 M UREA. Reduction and alkylation of captured proteins was performed by incubation with 10 mM dithiothreitol (DTT) for 60 min and by 20 mM iodoacetamide (IAA) in the dark for 30 min. The volume was then increased to 200 µL with 50 mM TEABC pH 8.0 containing 3 µg of sequencing grade trypsin. Overnight digestion carried out on a shaker maintained at 37°C. The supernatant was collected and the beads were washed once with 50 µL of 50 mM TEABC, which was collected and pooled. The final 450 µL was acidified to 1% trifluoroacetic acid, spun at 18,000 x *g* and was desalted with a homemade reversed-phase C_{18} column following stage-protocol previously published methods. In experiments designed to enrich biotinylated proteins using anti-biotin antibodies, after elution, the samples were dried and resuspended in 2 M Urea in 50 mM TEABC and digested with 3 µg of trypsin and desalted using a C_{18} column.

For BioSITe experiments the fractionated peptides were analyzed on an Orbitrap Fusion Lumos Tribrid Mass spectrometer coupled with the Easy-nLC 1200 nano-flow liquid chromatography system (Thermo Fisher Scientific). The peptides from each fraction were

reconstituted in 20 µL 0.1% formic acid and loaded onto an Acclaim PepMap 100 Nano-Trap Column (100 µm x 2 cm, Thermo Fisher Scientific) packed with 5 µm C_{18} particles at a flow rate of 4 µl per minute. Peptides were separated at 300-nl/min flow rate using a linear gradient of 7% to 30% solvent B (0.1% formic acid in 95% acetonitrile) over 95 min on an EASY-Spray column (50 cm x 75 µm ID, Thermo Fisher Scientific) packed with 2 µm C_{18} particles, which was fitted with an EASY-Spray ion source that was operated at a voltage of 2.3 kV.

Mass spectrometry analysis was carried out in a data-dependent manner with a full scan in the mass-to-charge ratio (m/z) range of 300-1,800 in the "Top Speed" setting, three seconds per cycle. MS and MS/MS were acquired for the precursor ion detection and peptide fragmentation ion detection, respectively. MS scans were measured at a resolution of 120,000 at an m/z of 200. MS/MS scans were acquired by fragmenting precursor ions using the higher-energy collisional dissociation (HCD) method and detected at a mass resolution of 30,000, at an m/z of 200. Automatic gain control for MS was set to one million ions and for MS/MS was set to 0.05 million ions. A maximum ion injection time was set to 50 ms for MS and 100 ms for MS/MS. MS was acquired in profile mode and MS/MS was acquired in centroid mode. Higher-energy collisional dissociation was set to 32 for MS/MS. Dynamic exclusion was set to 35 seconds, and singly-charged ions were rejected. Internal calibration was carried out using the lock mass option (m/z 445.1200025) from ambient air. Data acquisition of click chemistry-modified O-GlcNAc modified peptides were carried out using alternate HCD/ETD (electron transfer dissociation) method.

2.5 BioID and APEX preparation and lysis:

BioID was performed essentially as described Roux et al. with some minor modifications. Ten mg of protein per replicate was incubated with 200 µL of high capacity neutravidin for overnight at 4 ºC. After overnight incubation with 50 µM biotin, Ba/F3 cells expressing BirA*-BCR-ABL p190 were lysed in BioID lysis buffer (50 mM Tris, pH 7.5, 500 mM NaCl, 0.4 % SDS, 2 % Triton X-100 with Halt™ protease inhibitors) followed by sonication (3 rounds, duty cycle 30%, 20 second pulses). After cleaning, the samples were centrifuged at 16,000 x *g* for 10 min, equal volumes of 50 mM Tris were added to the BioID samples. Lysates were quantified by bicinchoninic acid (BCA) assay and 10 mg of protein per replicate was incubated with 200 µL of high capacity neutravidin overnight at 4 C. After incubation, the bead slurry was washed with 2% SDS, 50 mM Tris three times, BioID lysis buffer three times and 50 mM Tris three times and 50 mM triethylammonium bicarbonate (TEABC) three times. Cells used for experiments to enrich biotinylated proteins using anti-biotin antibodies were lysed in the same buffer used in the BioSITe strategy and the biotinylated proteins were captured and eluted in the same manner as described for BioSITe experiments. The experiments were carried out in triplicate using 10 mg of proteins per replicate.

APEX was performed as described by Hung et al. APEX-IMS and APEX-NES were transiently transfected into HEK293 using Lipofectamine 2000. One day later, the transfected cells were incubated with 50 µM biotin-phenol for 30 min and then with 1 mM H_2O_2 for 1 minute at room temperature. The APEX reaction was subsequently quenched by washing cells with quenching solution containing 10 mM sodium ascorbate, 5 mM Trolox and 10 mM sodium azide in PBS. Cells were collected by scrapping and pelleted.

2.6 Direct detection of biotin-containing tags:

This procedure was carried out as described by Schiapparelli et al. Ten mg of peptides generated from cells expressing BirA*-BCR-ABL p190 were dissolved in 1 mL of PBS and incubated with 90 µL of high capacity neutravidin beads for 1 hour at room temperature. Bead slurry was sequentially washed with PBS three times, 5 % acetonitrile three times and ultrapure water three times. Captured peptides were eluted by adding the elution buffer containing 0.2 % trifluoroacetic acid, 0.1 % formic acid, and 80 % acetonitrile. Peptides were eluted with and without boiling for 5 min for a total of 10 elutions which were pooled and desalted using a de-salting column as previously described.

2.7 BioSITe:

For each replicate, 10 mg of total protein from BirA*-BCR-ABL cells was digested into peptides and incubated with biotin-specific antibodies bound to protein G beads. For APEX samples, 8 mg of total protein was used per each replicate and digested into peptides. The bead slurry was sequentially thoroughly washed. Biotinylated peptides were eluted four times. Anti-biotin antibodies (100 µg) were coupled with 120 µL of protein G bead slurry for overnight at 4 °C. Antibody coupled beads were further washed with PBS once and BioSITe capture buffer (50 mM Tris, 150 mM NaCl, 0.5% Triton X-100) twice. Peptides were dissolved in 1 mL of BioSITe capture buffer. After dissolving peptides, pH was adjusted to neutral (7.0-7.5) and peptide BCA was preformed to estimate peptide concentration. Peptides were subsequently incubated with biotin antibody coupled protein G beads for 2 hours at 4 °C. The bead slurry was sequentially washed two times with BioSITe capture buffer, two times with 50 ml Tris and two times

with ultrapure water. Biotinylated peptides were eluted four times using elution buffer (80% acetonitrile and 0.2% trifluoroacetic acid in water). The eluent was further cleaned up using C_{18} reversed-phase column as previously described.

2.8 BioSITe for quantitative proteomics:

Peptides from the combined samples were subjected to BioSITe analysis. Cells expressing BirA*-BCR-ABL p190 or BirA*-BCR-ABL p210 were cultured overnight with 50 mM heavy or light biotin, respectively. The cells were lysed equal amounts from each condition were mixed to generate a total of 10 mg sample, which was digested, desalted, lyophilized and subjected to BioSITe as described above.

Biotinylation of O-GlcNAc sites using click-it chemistry- HEK293 cells were treated with 500 nm thiamet-G for 4 hrs to enhance the overall O-GlcNAcylation levels by inhibiting O-GlcNAcase (OGA), an enzyme that removes O-GlcNAc. Peptides were generated as previously described and in-solution Lys-C digestion was carried out overnight at 37°C. The peptide digest was then cleaned up using C18 solid phase extraction cartridge. The cleaned peptides were then vacuum dried and labeled with tandem mass tags. O-GlcNAc sites on these peptides were modified in-vitro using GalT1 enzymatic labeling kit. Briefly, peptides were reconstituted in 1.1 ml of 20 mM HEPES (pH 7.9), and 80 ul of 100 mM $MnCl_2$, 75 ul of 0.5 mM UDP-GalNAz, 62ul of GalT1 and 5 ul of PNGase F. The mixture was then incubated overnight at 4°C. The peptides were then reacted with 10uM DIBO Biotin-Alkyne in a copper-free Click-IT chemistry labeling of azides at room temperature for 1 hour. The peptides were then cleaned using C18 solid phase extraction cartridge followed by strong-cation exchange SPE and then vacuum dried. The

O-GlcNAc-modified and Click-IT labeled peptides were then subjected to BioSITe-based enrichment.

2.9 Mass spectrometry data analysis and post-processing and bioinformatics:

Peptides from on-bead digests of streptavidin captured biotinylated proteins, phosphotyrosine immunoprecipitation, and total protein lysates were analyzed using a nanoflow UPLC system (Easy nanoLC 1200, Thermo Scientific) interfaced with Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific) equipped with an EasySpray ion source. In total, 39 fractions from different BioID interactome enrichment experiments, anti-phosphotyrosine immunoprecipitation, and total protein experiments were analyzed by mass spectrometry. Three replicates of pTyr enrichment experiments, 12 bRPLC total protein fractions, 3 replicate streptavidin captures (run twice per replicate) comprised the experiments carried out in Ba/F3 cells. Similarly, 3 replicates of pTyr enrichments, 12 bRPLC total protein fractions, 6 replicates of label-free streptavidin captures comprised experiments in multi-potent progenitor (MPP) cells

Proteome Discoverer software suite (v 2.1; Thermo Fisher Scientific) was used for quantitation and database searches. The MS/MS data were searched using the SEQUEST search algorithm against a Mouse RefSeq database (v73 containing 58,485 entries) supplemented with frequently observed contaminants. Additionally, the BirA-p190$^{BCR\text{-}ABL}$ sequence, which represents the common sequence of p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ and the p210$^{BCR\text{-}ABL}$ specific sequence (DH and PH domain) were added to the database, as separate entries. The search parameters used for SEQUEST are listed in Supplemental Methods. The PSMs, peptides and proteins were filtered at 1% false discovery rate cut-off calculated using decoy database searches. After database searching, SILAC ratios of peptides were

quantified with PyQuant.[37] For label-free analysis of the MPP BioID experiment, MaxQuant software package with default settings was used.[38] Parameters used for SEQUEST search algorithm included trypsin as protease with full specific and a maximum of one allowed missed cleavage; carbamidomethylation of cysteine as a fixed modification; oxidation at methionine and biotinylation at Lysine as variable modification. In case of phosphotyrosine immunoprecipitation data analysis phosphorylation at tyrosine was selected as variable modification. For all the 3-State SILAC labeled samples, 13C6, 15N2-lysine, 2H4-lysine, 13C6-arginine and 13C6, 15N4-arginine as variable modifications. The precursor tolerance was set at 10 ppm while the fragment match tolerance was set to 0.1 Da.

In-house python scripts (https://github.com/pandeylab/pythomics) were used for post-processing annotation, protein inference, and normalization of the SILAC data. To correct for loading biases in the SILAC mixes in the pTyr experiments, we normalized the data based on the total protein measurements from 12 bRPLC fractions. Total protein was normalized to the median and that correction factor was applied to all values in pTyr data. After normalization, the total protein SILAC ratios for all peptides mapping to individual proteins were grouped and the median value was calculated for protein level measurements. For pTyr data, SILAC ratios were maintained at the peptide level and grouped by site, and median values were calculated, SILAC ratios for pTyr values were adjusted based on the corresponding abundance value from the total protein data that the phosphosite mapped. Additional median normalization was applied to the $p190^{BCR-ABL}/p210^{BCR-ABL}$ ratio as $p190^{BCR-ABL}$ cells were observed to have a slight increase in global phosphorylation compared to $p210^{BCR-ABL}$ cells in the Ba/F3 experiment, the opposite was

observed in MPP experiment and was normalization similarly. Finally, when a phosphorylation site was sampled in all three replicates, the median was calculated, in the case of phosphorylation sites sampled in only two replicates, the average was taken. The SILAC BioID data followed a similar scheme although based on the ratios of known endogenously biotinylated proteins being near equally represented in the $p190^{BCR-ABL}$/$p210^{BCR-ABL}$ cells, no normalization was applied to the data. Median values of replicates were calculated when sampled in all three replicates and average was used for sampling in 2 replicates. No correction was applied for abundance values differences from the total protein. BioID data went through a series of filtering steps to eliminate potential false-positive interactors. Specifically, we first compared all putative interacting proteins (irrespective of BCR-ABL variant) to a negative control streptavidin capture consisting of non-BirA Ba/F3 lysate to screen for endogenously biotinylated proteins. Next, we referenced the CRAPome database[39] to identify and remove proteins with a higher likelihood of being non-specific in pull-down assays, to further filter potential non-specific protein interactors. Then proteins were filtered if they were not sampled in at least 2 of the replicates. We followed these initial filtering steps with STRING analysis,[40] a tool to search interaction networks between proteins based on previous studies, on the remaining protein identities to identify those proteins with known interactions by biochemical evidence. This analysis revealed a major node comprising of many known BCR-ABL interactors, indicating the robustness of our assay in detecting these interactions. Our final set of protein interactors after the filtering was comprised of 90 proteins. For BioSITe experiments Proteome Discoverer (v 2.1; Thermo Scientific) suite was used for quantitation and identification using all 3 replicate LC-MS/MS runs

per experiment searched together. Spectrum selector was used to import spectrum from raw file. During MS/MS preprocessing, the top 10 peaks in each window of 100 m/z were selected for database search. The tandem mass spectrometry data were then searched using SEQUEST algorithm against protein databases (For BioID experiments; mouse NCBI RefSeq 73 (58,039 entries) with the addition of fasta file entries for BCR-ABL p190 and the DH and PH domain of BCR-ABL p210, for APEX experiments; human NCBI RefSeq (73,198 entries) with the additions of fasta file entries of IMS-APEX2 and NES-APEX2 constructs) with common contaminant proteins. The search parameters for identification of biotinylated peptides were as follows: a) trypsin as a proteolytic enzyme (with up to three missed cleavages); b) peptide mass error tolerance of 10 ppm; c) fragment mass error tolerance of 0.02 Da; and d) carbamidomethylation of cysteine (+57.02146 Da) as a fixed modification and oxidation of methionine (+15.99492 Da) and biotinylation of lysine (+226.07759 Da) as variable modifications. The search parameters for the identification of biotin-phenol modified peptides were as follows: a) trypsin as a proteolytic enzyme (with up to two missed cleavages); b) peptide mass error tolerance of 10 ppm; c) fragment mass error tolerance of 0.02 Da; and d) carbamidomethylation of cysteine (+57.02146 Da) as a fixed modification and oxidation of methionine (+15.99492 Da). Biotinylation of lysine (+226.07759 Da), biotin-phenol modification of tyrosine (+361.14601 Da), and oxidized-biotin-phenol modification of tyrosine (+377.141 Da) were all used as variable modifications. For the identification and quantification of the peptides modified by light or heavy biotin, all the raw files from the three replicates were searched together. The search parameters for identification of either light or heavy biotinylated peptides were as follows: a) trypsin as a proteolytic enzyme

(with up to three missed cleavages); b) peptide mass error tolerance of 10 ppm; c) fragment mass error tolerance of 0.02 Da; and d) carbamidomethylation of cysteine (+57.02146 Da) as a fixed modification and oxidation of methionine (+15.99492 Da), light biotinylation of lysine (+226.07759 Da) and heavy biotinylation of lysine (+230.103 Da) as variable modifications. The minimum peptide length was set to 6 amino acids. For identification of click chemistry-modified O-GlcNAcylated peptides, apart from Oxidation of methionine and Carbamidomethylation of Cysteine, variable modification of click label (993.36 Da i.e. HexNAc+GalNAz+DIBO alkyne biotin) on serine and threonine residues were included in database search. Peptides and proteins were filtered at a 1% false-discovery rate (FDR) at the PSM level using percolator node and at the protein level using protein FDR validator node, respectively.

The protein quantification was performed with following parameters and methods. The heavy to light ratios of the biotinylated peptides were measured by the Precursor Ions Quantifier node. Unique and razor peptides both were used for peptide quantification while protein groups were considered for peptide uniqueness. Precursor ion abundance was computed based on intensity and the missing intensity values were replaced with the minimum value. Protein grouping was performed with strict parsimony principle to generate the final protein groups. All proteins sharing the same set or subset of identified peptides were grouped while protein groups with no unique peptides were filtered out. The Proteome Discoverer iterated through all spectra and selected PSM with the highest number of unambiguous and unique peptides.

Identified protein and peptide spectral match (PSM) level data were exported as tabular files from Proteome Discoverer 2.1. We used an in-house Python script to compile the

peptide level site information mapped to UniProt or RefSeq databases. The summary count on number of supported peptides and PSMs are then calculated at protein level. In case of the heavy-biotin experiment, the ratios are calculated from total channel intensities of identified peptides.

2.10 Protein modeling:

The degree and location of biotinylation as well as the protein domain organization were assigned based on information in UniProt (uniprot.org) taking into account the protein size, number of total lysines and domain annotation according to PROSITE and InterPro databases. To map biotinylation sites on three-dimensional structures, structural models were obtained from the Protein Data Bank (PDB, rcsb.org) as far as available or homology models were created using Swiss-Model (swissmodel.expasy.org).

2.11 Data availability:

All mass spectrometry data and search results have been deposited to the ProteomeXchange Consortium) via the PRIDE partner repository with the dataset identifier PXS005209 and PXD007862.

2.12 Experimental Design and Statistical Rationale:

All experiments described in this study were performed as three independent replicates from the same starting material (process replicates). Because this is a new method, this replicate design was employed to account for sampling bias. A qualitative comparison of the identifications from the methods being compared was done in all cases except in the experiment involving heavy biotin where the median fold-ratios were plotted.

CHAPTER 3.

DIFFERENTIAL SIGNALING THROUGH BCR-ABL VARIANT FUSION PROTEINS

The following work was published in the journal Leukemia and all reference to supplemental data/figures can be found at the Leukemia website with the following this citation[41].

3.1 Experimental rationale

Both $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ are constitutively active kinases with similar molecular makeup that differ only by the presence ($p210^{BCR\text{-}ABL}$) or absence ($p190^{BCR\text{-}ABL}$) of a DH and PH domain (Figure 1a). Given the differences in clinical outcomes, we hypothesized that these two proteins display differential binding of partner proteins and phosphorylation signatures leading to distinct downstream signaling pathways. We therefore employed a two-pronged approach by expressing the variants in the same cellular background, the well-studied murine hematopoietic Ba/F3 cell line. The first part of our approach relied on detecting differential protein-protein interactions ("interactome") of these two proteins by using BioID coupled with mass spectrometry and SILAC-based quantitation (Figure 1b). The second part involved investigation of differential global phosphorylation signatures ("phosphoproteome") via pTyr peptide enrichment followed by mass spectrometry and SILAC-based quantitation[42] (Figure 1c).

3.2 Establishment of BioID system to examine BCR-ABL interactome in Ba/F3 cells

BioID can detect both transient interactions as well as more stable interactions. When the promiscuous biotin ligase (BirA) is cloned in-frame with a protein of interest, it biotinylates interacting proteins including those that are transiently bound, enabling detection of interacting proteins via enrichment by streptavidin.[43] We cloned BirA onto the N-terminus of both $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ because we reasoned that this positions the BirA moiety most proximal to differing domains of $p190^{BCR\text{-}ABL}$ and

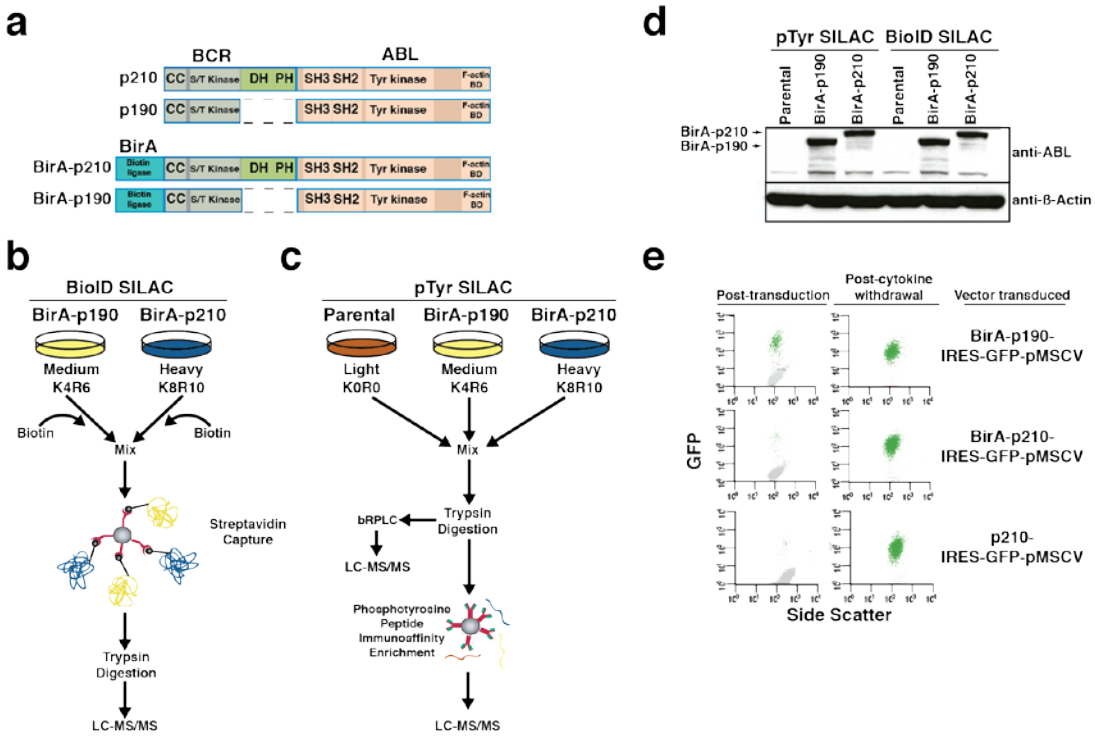

Figure 3.1 Strategy for investigating BCR-ABL interactome and phosphoproteome. (a) Domain structures of $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ are shown. The lack of Dbl homology (DH) and Pleckstrin homology (PH) domain in the BCR part of the $p190^{BCR\text{-}ABL}$ variant is indicated by dashed lines. The biotin ligase, BirA, was cloned in-frame at the N-terminus of both variants as indicated. The domains shared by the two variants are coiled-coiled (CC), BCR serine/threonine Kinase (S/T Kinase), Src homology 3 (SH3), Src Homology 2 (SH2) Abl Tyrosine Kinase (TyrK), and F-Actin Binding Domain (F-Actin BD). (b) The experimental workflow for interactome analysis by BioID in Ba/F3 cells. Both BCR-ABL cell lines were grown in SILAC containing K4R6(medium) or K8R10 (heavy) amino acids, exogenous biotin was added to both cultures as indicated and incubated for 24 hours. The resulting lysates were subjected to streptavidin capture followed by trypsin digestion and LC-MS/MS. (c) The experimental workflow for phosphoproteome analysis in Ba/F3 cells. Cells were grown in SILAC media containing K0R0 (light), K4R6 (medium), or K8R10 (heavy) amino acids. The parental line was cytokine starved and the cultures harvested, lysed, mixed, and trypsin digested. A small fraction of the lysate was fractionated by reversed phase liquid chromatography (bRPLC) and analyzed by LC-MS/MS. The remaining peptides were immunoprecipitated with pY1000 antibody and subjected to LC-MS/MS analysis. (d) Western Blot analysis with antibodies against Abl and β-Actin showing the relative expression levels of the two variants in Ba/F3 cells. (e) Addition of BirA does not alter the ability of BCR-ABL to drive cytokine independence in Ba/F3 cells: Flow cytometry plots monitoring GFP expression versus side scatter in Ba/F3 cells expressing BCR-ABL isoforms with or without BirA after retroviral transduction and post IL-3 withdrawal are shown.

p210$^{BCR\text{-}ABL}$ (Figure 1a). Relative expression levels of BirA tagged p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ proteins were analyzed by immunoblot (Figure 1d). In agreement with previous studies in which expression of BCR-ABL in Ba/F3 cells induced IL-3 independence,[44] the BirA tagged versions of both p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ also induced cytokine independence (Figure 3.1e). In order to detect interacting proteins, these cells were cultured in the presence of exogenous biotin to induce *in vivo* labeling of BCR-ABL variant interactors. These cells were then lysed, mixed and three independent experiments were performed to enrich biotinylated proteins by streptavidin capture under denaturing conditions, followed by trypsin digestion. The resulting peptide mixture was then subjected to LC-MS/MS analysis (Figure 3.1b). As expected, a substantial enrichment of biotinylated proteins as well as the capture of BCR-ABL variants was observed in the streptavidin isolates (Supplemental Figure 1a). A total of 814 proteins were identified by LC-MS/MS analysis, of which 308 proteins were detected in all three replicates and their corresponding SILAC ratios are plotted in Figure 3.2a (Supplemental Table 1 for complete data). Additional streptavidin isolates were generated for immunoblot based validation of targets identified in the mass spectrometry data (Figure 3.2b).

3.3 A common BCR-ABL interactome signature

First, we identified the overall BCR-ABL interactome within our BioID experiment by employing filtering steps to remove non-specific interactions and identify well supported interactions using literature-based tools (see Materials and Methods). This analysis resulted in a subset of 90 proteins, whose SILAC ratios are represented in

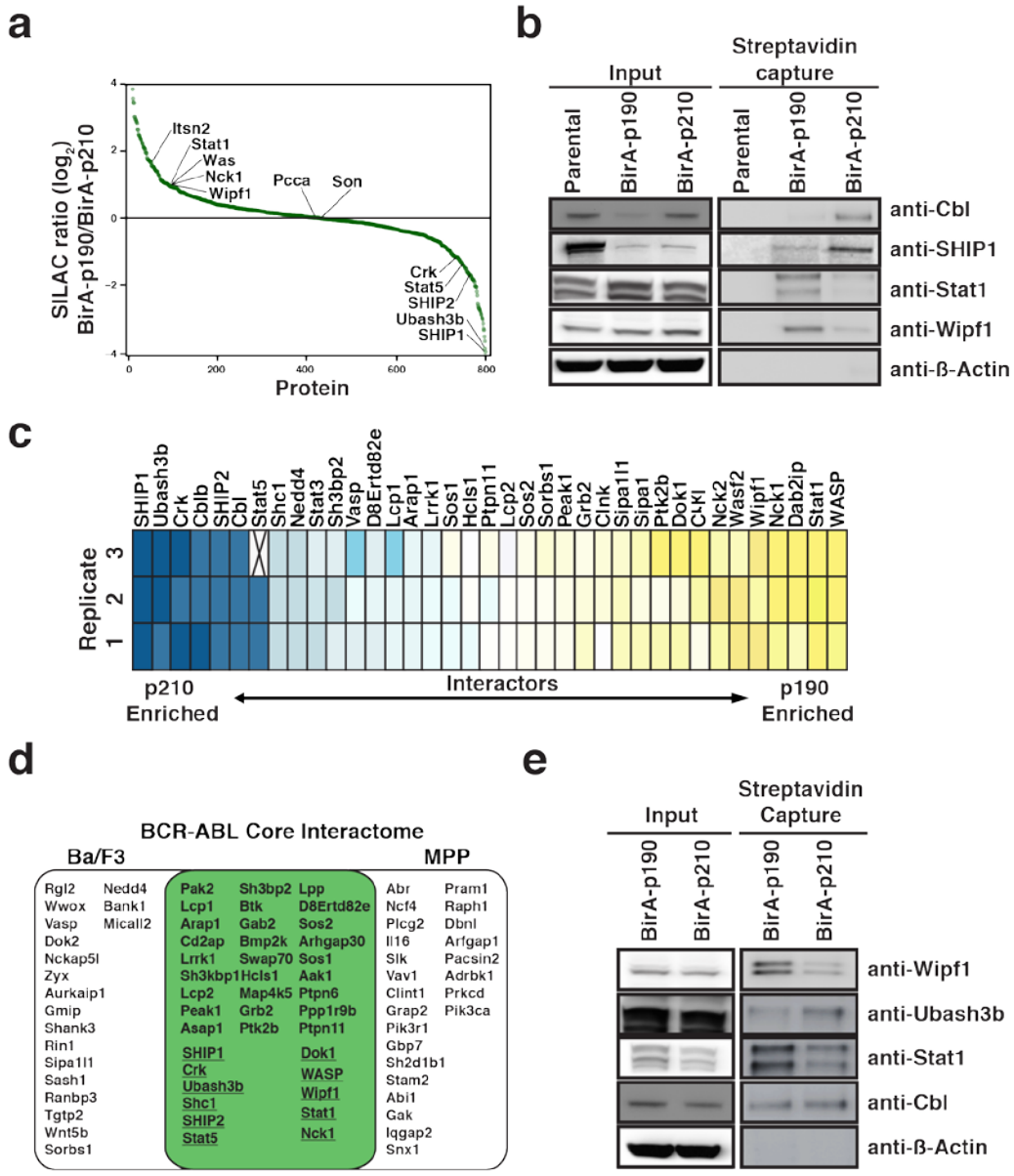

Figure 3.2 Interactome analysis of BCR-ABL variants using the BioID system. (a) A plot of SILAC ratios (log_2 scale) for the detected proteins in Ba/F3 interactome analysis. Endogenously biotinylated proteins are shown. (b) Western blot validation of Ba/F3 mass spectrometry BioID data. Biotinylated proteins captured by streptavidin from the indicated cells were probed with antibodies against Cbl, Inpp5d/SHIP1, Stat1, Was/WASP, and β-Actin as shown in the right panel. The whole cell lysate is also probed in parallel (left panel). (c) Heat map of select BCR-ABL interacting proteins along with any observed enrichment in one or the other BCR-ABL isoform is shown. (d) Venn diagram of the BCR-ABL core interactome observed in our analysis of Ba/F3 and MPP cells. Within the green are overlapping interactors that were detected in both MPP and Ba/F3 cells. Interactors in white are either cell type specific or not sampled in the cell type. Gene symbols for overlapping $p190^{BCR\text{-}ABL}$ enriched and $p210^{BCR\text{-}ABL}$ enriched interactions are underlined. (e) Western blot validation of MPP mass spectrometry BioID data. Biotinylated proteins captured by streptavidin from the indicated cells were probed with antibodies against Wipf1, Ubash3b, Stat1, Cbl, and β-Actin as shown in the right panel. The whole cell lysate is also probed in parallel (left panel).

the heat map in Figure 2c. This subset includes many known BCR-ABL interactors including Nck adaptor protein 1 (Nck1), Nck adaptor protein 2 (Nck2), Docking protein 1 (Dok1), Tyrosine-protein phosphatase non-receptor type 11 (Ptpn11), Son of sevenless homolog 1 and 2 (Sos1, Sos2), Abl interactor 1 (Abi1), Grb2-associated-binding protein 2 (Gab2), and Pseudopodium-enriched atypical kinase (Peak1), among others. We also confirmed previously defined components of the BCR-ABL interactome[25] in our analysis: Growth factor receptor-bound protein 2 (Grb2), SHC-transforming protein 1 (Shc1), CRK proto-oncogene adaptor protein (Crk), Cbl proto-oncogene (Cbl), phosphoinositide-3-kinase regulatory subunit 2, Ubash3b, and Inositol Polyphosphate Phosphatase Like 1 (Inppl1 or SHIP2) (Figure 3.2d and Supplemental Table 1).

3.4 $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ display differential interactions with proteins in different cellular compartments

In order to characterize the *differential* interactome of $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$, we measured SILAC ratios of peptides from interacting biotinylated proteins captured from the two variants. We defined $p190^{BCR\text{-}ABL}$ enriched interactions as proteins with SILAC ratios indicating greater than 1.8-fold enrichment in $p190^{BCR\text{-}ABL}$, while ratios indicating greater than 1.8-fold enriched in $p210^{BCR\text{-}ABL}$ were defined as $p210^{BCR\text{-}ABL}$ enriched interactors. We also defined those proteins displaying ratios that indicated greater than 1.5-fold but less than 1.8-fold difference between the cell types as trending towards an enrichment of interaction. As a control we monitored the SILAC ratios of known endogenously biotinylated proteins commonly found in BioID experiments[43] and those were not enriched in either cell type (Figure 3.2a). We then examined the differential interactors within the subset of 90 proteins identified after filtering steps.

Many of the molecules identified to be enriched in their interactions with $p190^{BCR\text{-}ABL}$ are cytoskeletal proteins. Members of the Wiskott-Aldrich syndrome protein family, specifically, WASP, WAS/WASL-interacting protein family member 1 (Wipf1) and 2 (Wipf2), and WAS Protein Family Member 2 (Wasf2) all displayed >1.8-fold increased association with $p190^{BCR\text{-}ABL}$ relative to $p210^{BCR\text{-}ABL}$. Interestingly, BCR-ABL has been previously shown to interact with the cytoskeleton and has been implicated in inducing cytoskeletal-related phenotypes, although these alterations were not assigned to a specific variant[45–49]. The differential interaction with Wipf1 identified through our quantitative mass spectrometry data was confirmed by immunoblot analysis of streptavidin enriched isolates from $p190^{BCR\text{-}ABL}$, $p210^{BCR\text{-}ABL}$ and parental (as a negative control) cell lysates (Figure 3.2b). Adaptor proteins, such as Nck1, Nck2, and intersectin 2, also known as SH3p18-Like WASP-Associated Protein (Itsn2), are not explicitly cytoskeletal but have been shown to mediate cytoskeleton signaling pathways, also displayed an increased interaction with $p190^{BCR\text{-}ABL}$ in our analysis. Finally, signal-induced proliferation-associated protein 1 (Sipa1) and signal-induced proliferation-associated 1-like protein 1 (Sipa1l1), also modulators of the cytoskeleton, exhibited a modest preferential interaction with $p190^{BCR\text{-}ABL}$. Sipa1 has been previously shown to interact with BCR-ABL and displays co-localization to uropod actin structures of migrating cells, suggesting that BCR-ABL is likely found at these structures.[50] Here, we show that these important cytoskeletal re-modeling factors preferentially interact with the $p190^{BCR\text{-}ABL}$ variant.
In contrast to the findings for $p190^{BCR\text{-}ABL}$, many of the proteins that displayed increased association with $p210^{BCR\text{-}ABL}$ (Figure 3.2c) are plasma membrane proximal proteins or display an increased activity at the plasma membrane. For example, Ubash3b, a

membrane proximal adapter protein with phosphatase activity that targets Src and Syk family kinases and regulates plasma membrane bound receptor tyrosine kinases,[51,52] was found to be 15-fold enriched in its interaction with $p210^{BCR\text{-}ABL}$ compared to $p190^{BCR\text{-}ABL}$. Ubash3b is also a known interactor and negative regulator of Cbl[51] and, intriguingly, Cbl and its homolog Casitas B-lineage lymphoma b (Cblb), were both enriched for $p210^{BCR\text{-}ABL}$ interaction in our analyses (7-fold and 4-fold, respectively). Cbl has previously been shown to form a complex with BCR-ABL and mediate the degradation of SHIP1 via polyubiquitination.[53,54] Our data confirm the downregulation of SHIP1 upon expression of BCR-ABL, (compare inputs of both variant expressing cells to the parental line in Figure 2b). Moreover, in addition to SHIP1, total protein levels of Cbl and Cblb are also downregulated compared to parental cells. Interestingly, the levels of SHIP1/Cbl/Cblb are downregulated in $p190^{BCR\text{-}ABL}$ to a greater extent than $p210^{BCR\text{-}ABL}$ suggesting that both forms are engaged in this complex but $p190^{BCR\text{-}ABL}$ might be mediating increased degradation of SHIP1 compared to $p210^{BCR\text{-}ABL}$. The preferential interaction of Cbl and SHIP1 with $p210^{BCR\text{-}ABL}$ along with assessment of their protein abundance was confirmed by immunoblotting streptavidin isolates and whole cell lysates (Figure 3.2b). A combination of higher total protein levels of SHIP1 in $p210^{BCR\text{-}ABL}$ cells and the presence of the $p210^{BCR\text{-}ABL}$ enriched interaction of Ubash3b could also play an inhibitory role in the degradation SHIP1. Ubash3b is a ubiquitin binding protein and could be binding to ubiquitinated SHIP1, thus attenuating its degradation. Other membrane proximal proteins identified to be enriched in interaction with $p210^{BCR\text{-}ABL}$ include the phosphatase SHIP2 and adaptor proteins Crk and Shc1.

Finally, our data suggest that p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ engage with Stat proteins differently. For example, Stat1 showed a 2-fold enrichment in p190$^{BCR\text{-}ABL}$-specific interactions, which was confirmed by immunoblotting (Figure 3.2b). In addition, both Stat5 and Stat3 were found to be enriched (2.5-fold and 1.5-fold) in p210$^{BCR\text{-}ABL}$ mediated interactions. Importantly, previous studies have indicated that Stat5 is a direct substrate of BCR-ABL.[55]

3.5 p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ interaction analysis in MPP cells

In an effort to confirm our findings, we interrogated p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ interaction differences in an alternate cell type. We chose an *ex vivo* expanded primary murine hematopoietic progenitor cell derived from early B cell factor 1 (Ebf1) null fetal livers. These cells exhibit a "multi-potent progenitor" phenotype and have been shown to be an excellent model for hematopoiesis prior to B cell specification since these cells can be induced in culture to differentiate toward the B, T and myeloid lineage.[35,36,56,57] The p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ interaction analysis in MPP cells is diagramed in Supplemental Figure 2a. Interrogating the MPP cells expanded our experimentally derived BCR-ABL interactome and the overlap with the Ba/F3 experiment is shown in Figure 3.2d. Many of the same molecules were identified in the interactome analysis in both cell types and these analyses also uncovered some potentially cell-type specific interactions (Figure 3.2d). Importantly, many of the specific differential interactions observed in the Ba/F3 cells were also observed in these cells. For example, the preferential interaction of Dok1, WASP, Wipf1, Stat1, and Nck1 with p190$^{BCR\text{-}ABL}$ found in the MPP cells was in agreement with the Ba/F3 experiment, although the enrichment for these molecules was not as pronounced as was observed in Ba/F3 cells, perhaps due

to cell-type specific differences or due to the sensitivity of the label-free quantitation used in MPP experiments. In order to confirm these interactions, we measured Wipf1, Stat1, and Ubash3b levels by subjecting the streptavidin isolates to immunoblotting and these data highlighted a clear p190$^{BCR\text{-}ABL}$ skewed interaction in MPP cells (Figure 3.2e). Cbl and Cblb, which had shown robust increased interactions with p210$^{BCR\text{-}ABL}$ relative to p190$^{BCR\text{-}ABL}$ in Ba/F3 cells, were found to also display specific differential interactions in the MPP experiment, although, again, with more modest enrichments (Figure 3.2e and Supplemental Table 2). As in the experiments in the Ba/F3 system, p210$^{BCR\text{-}ABL}$ expressing MPPs displayed increased interactions with SHIP1, Ubash3b, SHIP2, Stat5, and Shc1 relative to p190$^{BCR\text{-}ABL}$.

3.6 Global phosphorylation signatures of p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ in Ba/F3 cells

It is possible that increased interaction of the p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ variants with a specific set of proteins leads to the increased phosphorylation of those targeted proteins to induce unique signaling through downstream pathways. To identify and quantify phosphotyrosine sites that are induced or altered in either p190$^{BCR\text{-}ABL}$ or p210$^{BCR\text{-}ABL}$ harboring cells, we employed a 3-state SILAC strategy. Specifically, we isolated lysates from Ba/F3 cells harboring BirA-tagged versions of p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ (as used in the BioID experiments) and, as a control for basal phosphorylation levels, from Ba/F3 cells containing no exogenous protein. Not surprisingly, global phosphorylation was observed to be higher in cells ectopically expressing the constitutively active kinases, p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ relative to the parental cells. This was observed when lysates were enriched for pTyr containing proteins by immunoprecipitation with a pTyr specific antibody (Figure 3.3a). Three

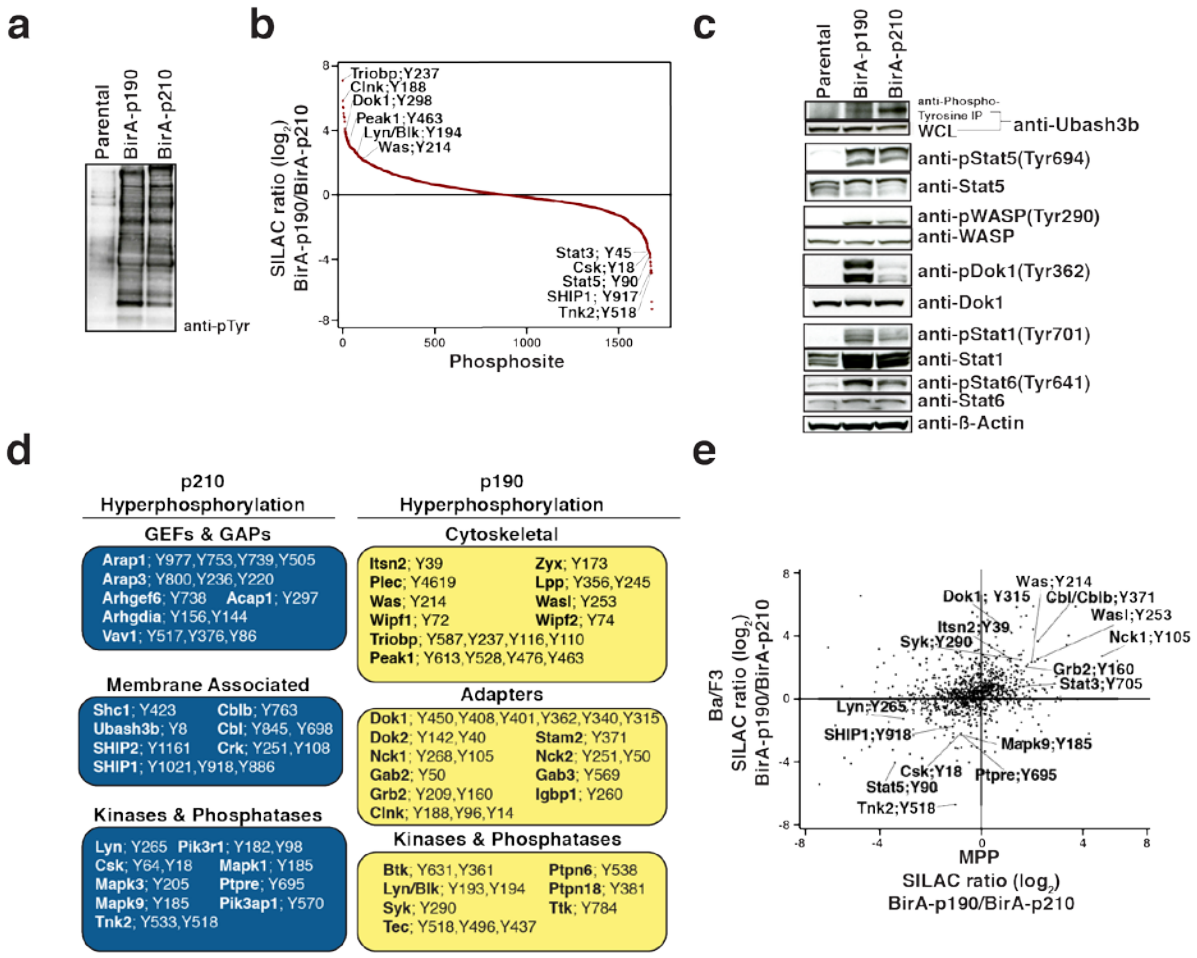

Figure 3.3 Tyrosine phosphorylation analysis of BCR-ABL variants. (a) Tyrosine phosphorylated proteins were immunoprecipitated from whole cell lysates (WCL) of parental, $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ Ba/F3 cells using anti-phosphotyrosine antibody and probed with the same antibody as indicated. (b) Relative abundance of normalized tyrosine phosphopeptides based on their SILAC ratios (log_2 scale) in $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ expressing Ba/F3 cells. Selected proteins are labeled by name and corresponding phosphosite. (c) Validation of tyrosine phosphorylation by western blot analysis in Ba/F3 cells. All antibodies used were phosphosite specific with the exception of Ubash3b (in which case anti-phosphotyrosine antibody was used for the IP). Antibodies against the proteins were used to detect the overall protein expression in each case. (d) A list of notable tyrosine containing peptides hyperphosphorylated in either $p190^{BCR\text{-}ABL}$ or $p210^{BCR\text{-}ABL}$ Ba/F3 cells. The sites are grouped into categories of protein function; guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs), cytoskeletal, membrane associated proteins, adapters, and kinases/phosphatases. (e) Plot of SILAC ratio values of MPP the $p190^{BCR\text{-}ABL}/p210^{BCR\text{-}ABL}$ versus Ba/F3 the $p190^{BCR\text{-}ABL}/p210^{BCR\text{-}ABL}$. Selected proteins are labeled by name and corresponding phosphosite.

technical replicates of pTyr enrichments were subjected to LC-MS/MS and subsequent analyses yielded quantitation of 1,568 pTyr sites on 965 proteins, 746 sites of which were sampled in all three replicates. We measured phosphorylation ratios by calculating the median of replicate SILAC ratio $p190^{BCR\text{-}ABL}/p210^{BCR\text{-}ABL}$ values of sites (all SILAC ratios are plotted in Figure 3.3b and all data is provided in Supplemental Table 3). Phosphosites displaying ratio values greater than 2-fold enriched in $p190^{BCR\text{-}ABL}$ were defined as $p190^{BCR\text{-}ABL}$ hyperphosphorylated, while values with 2-fold enriched in $p210^{BCR\text{-}ABL}$ were defined as $p210^{BCR\text{-}ABL}$ hyperphosphorylated. We also define those sites displaying ratios that indicate greater than 1.5-fold but less that 2-fold difference between the cell types as trending toward hyperphosphorylation. For a select number of phosphorylation sites, the mass spectrometry results were confirmed by immunoblotting with tyrosine phosphorylation specific antibodies (Figure 3.3c).

3.7 Cytoskeleton proteins hyperphosphorylated in $p190^{BCR\text{-}ABL}$ expressing cells

Strikingly, in very strong agreement with our analysis of $p190^{BCR\text{-}ABL}$-skewed protein-protein interactions, we identified many sites on cytoskeletal proteins and cytoskeletal signal transduction pathway members to be hyperphosphorylated in response to $p190^{BCR\text{-}ABL}$ expression (Figure 3.3d). The hyperphosphorylation of the Wiskott-Aldrich syndrome family member proteins WASP, its homolog Neural Wiskott-Aldrich syndrome protein (Wasl or N-WASP), and their interacting proteins Wipf1/2 are of particular interest as we also observed WASP and Wifp1 to be enriched in $p190^{BCR\text{-}ABL}$ interaction. This combination of interaction and hyperphosphorylation not only implicates Wiskott-Aldrich syndrome family members as direct substrates of $p190^{BCR\text{-}ABL}$ but also further suggests that cytoskeletal pathways are being activated differently in

p190$^{BCR\text{-}ABL}$ expressing cells compare to p210$^{BCR\text{-}ABL}$, as these proteins are potent effectors of these pathways.[58] The specific function of the sites on the Wiskott-Aldrich syndrome family members observed to be p190$^{BCR\text{-}ABL}$ hyperphosphorylated further support the activation of this pathway. For example, sites Tyr293 and Tyr253 on WASP and N-WASP, respectively, are essential for the activation of both molecules.[58] WASP-Tyr293 hyperphosphorylation was confirmed by immunoblotting with a phospho-specific antibody and shown in Figure 3c. Activated N-WASP is essential for actin assembly and, via recruitment of Nck1 and Grb2, promotes actin based movements.[59] In agreement with the previous finding, we observed hyperphosphorylation of many sites on Nck1, Nck2, and Grb2 in p190$^{BCR\text{-}ABL}$ expressing cells. Both Nck1 and Nck2 were also identified as preferentially interacting with p190$^{BCR\text{-}ABL}$, suggesting that this BCR-ABL variant is co-located with all these proteins. The protein complex of Nck1/Grb2/Wiskott-Aldrich syndrome family members has been shown to be mediated by Abl-dependent phosphorylation of Tyr361 on Dok1. Phosphorylation of Dok1 recruits Nck1 to interact with Wiskott-Aldrich syndrome family members which initiates actin polymerization and filopodia in mouse embryonic fibroblasts.[60] Remarkably, we identified 9 tyrosine sites, including Tyr361, on Dok1 that are hyperphosphorylated in p190$^{BCR\text{-}ABL}$ expressing cells, further implicating enriched association of p190$^{BCR\text{-}ABL}$ with the cytoskeleton compared to p210$^{BCR\text{-}ABL}$. Importantly, Dok1 trended towards an increase of p190$^{BCR\text{-}ABL}$ interaction in the interactome data. Taken together, the combination of preferential interaction and hyperphosphorylation of key residues of Wiskott-Aldrich syndrome family members, Nck1/2, Grb2, and Dok1 suggest that p190$^{BCR\text{-}ABL}$ is profoundly influencing the cytoskeleton.

3.8 Phosphorylation signatures of p210$^{BCR-ABL}$ expressing cells

p210$^{BCR-ABL}$ displayed preferential protein-protein interactions with some important membrane-proximal proteins, including Ubash3b, SHIP2, Cbl, Cblb, and SHIP1. In agreement with this p190$^{BCR-ABL}$/p210$^{BCR-ABL}$ differential interaction, we also observed differential phosphorylation signatures indicating that the variants are exerting increased influence on these proteins and pathways. In general, increased p210$^{BCR-ABL}$ interaction was coupled with increased phosphorylation. For example, a novel phosphotyrosine site on Ubash3b (Tyr9) was observed to be hyperphosphorylated in p210$^{BCR-ABL}$ cells. To validate this novel site, we used an anti-pTyr antibody to isolate all pTyr proteins in the cells, and followed this immunoprecipitation with immunoblotting for Ubash3b proteins (Figure 3.3c). A direct comparison of total protein versus pTyr enriched, shows that Ubash3b is indeed more highly phosphorylated in p210$^{BCR-ABL}$ expressing cells (compare total protein with anti-pTyr enriched) leading us to conclude that Ubash3b is more active in p210$^{BCR-ABL}$ expressing cells. Other examples of concurrent p210$^{BCR-ABL}$ interaction and hyperphosphorylation were found in SHIP1 and SHIP2. As for SHIP1 phosphorylation we observed p210$^{BCR-ABL}$ hyperphosphorylation on the sites Tyr917 and Tyr1020. The SHIP1 sites Tyr917 and Tyr1020 are interesting because Dok1/Shc1 compete for binding at Tyr1020.[54,61,62] This might be noteworthy because, in our BioID analysis, Shc1 is favored by p210$^{BCR-ABL}$ and Dok1 by p190$^{BCR-ABL}$, implicating some interesting differential interplay between Shc1/Dok1/SHIP1 with respect to the two BCR-ABL variants. SHIP2 showed a robust increase in phosphorylation on Tyr1161 (2.4-fold enrichment over p190$^{BCR-ABL}$ expressing cells) and with additional phosphosites Tyr986, Tyr835, and Tyr662 all trending toward enrichment

(>1.5-fold) in p210$^{BCR\text{-}ABL}$ expressing cells. The site Tyr1136 (Tyr1135 in human) on SHIP2 showed a reciprocal relationship between interaction and phosphorylation as it was observed to be hyperphosphorylated in p190$^{BCR\text{-}ABL}$ expressing cells. Phosphorylation of this site has previously been shown to translocate SHIP2 to focal adhesions[63] which further implicates a connection of p190$^{BCR\text{-}ABL}$ to the cytoskeleton and suggests that the two variants use SHIP2 in different ways to promote their unique signaling programs. Cbl and Cblb represent another example against the general trend of increased interaction leading to increased phosphorylation. Specifically, Cbl and Cblb display a shared phosphopeptide that is hyperphosphorylated in p190$^{BCR\text{-}ABL}$ cells (where the interaction is less). These sites, Tyr369/363 (Tyr371/363 in human) increase the ubiquitin ligase activity of Cbl when phosphorylated.[64] This observation has particularly intriguing implications for poly-ubiquitination/degradation of SHIP1 by Cbl/BCR-ABL and could explain in part why SHIP1 is less abundant in p190$^{BCR\text{-}ABL}$ expressing cells due to increased degradation.

The rhoGEF domain in p210$^{BCR\text{-}ABL}$ has been shown to be constitutively activated and mutations that disrupt the activity of this domain decrease the transforming potential of the oncogene.[65] We wanted to investigate if any G-protein signaling pathway related proteins were differentially regulated by phosphorylation, which could identify those potentially working in concert with p210$^{BCR\text{-}ABL}$ to produce specific rhoGEF pathway signaling programs. We did observe many tyrosine sites on guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs) that were hyperphosphorylated in p210$^{BCR\text{-}ABL}$ cells (Figure 3.3d). For example, Arf-GAP with a rho-GAP domain, ANK repeat and PH domain-containing protein 1 and 3 (Arap1/CENTD2 and

Arap3/CENTD3), Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 1 (Acap1), rho guanine nucleotide exchange factor 6 (Arhgef6), rho GDP-dissociation inhibitor 1 (Arhgdia), and proto-oncogene Vav 1 (Vav1) are all hyperphosphorylated.

3.9 Differential regulation of Stat proteins

Previous work using immunoprecipitation and immunoblotting in the Ba/F3 system has revealed that Stat family members are differentially regulated by $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$[12,21]. In one study, Stat1 and Stat6 were found to show higher phosphorylation in $p190^{BCR\text{-}ABL}$ expressing Ba/F3 cells, while Stat5 phosphorylation was higher in $p210^{BCR\text{-}ABL}$ expressing Ba/F3 cells and Stat3 exhibited no change in phosphorylation levels. To confirm and extend this observation, which was limited by available antibodies and only detected overall levels of phosphorylation, we interrogated our differential pTyr phosphorylation data to determine differences at specific sites of phosphorylation induced by the BCR-ABL variants (summarized in Table 1). We identified a total of nine sites on Stat family proteins (Stat1, Stat2, Stat3, Stat5, and Stat6), six of which displayed hyperphosphorylation in either $p190^{BCR\text{-}ABL}$ or $p210^{BCR\text{-}ABL}$ expressing cells. We confirmed several of these sites, identified in our mass spectrometry data, by immunoblotting extracts from $p190^{BCR\text{-}ABL}$ or $p210^{BCR\text{-}ABL}$ expressing cells using commercially available phosphorylation site-specific antibodies (Figure 3.3c). These data confirm that sites Tyr701 in Stat1 and Tyr641 in Stat6 are hyperphosphorylated preferentially in $p190^{BCR\text{-}ABL}$ cells. Phosphorylation of Tyr701 and Tyr641, located just outside the SH2 domain, is required for dimerization, nuclear translocation, and DNA binding activity of both Stat1 and Stat6.[66,67] A similar site

Tyr694, as well as neighboring sites Tyr668 and Tyr682/3 in the SH2 domain, in Stat5a/b showed no change by mass spectrometry or immunoblotting. However, in p210$^{BCR\text{-}ABL}$ expressing cells, we identified an increase of phosphorylation at Tyr90 and Tyr114, which is located near the DNA binding domain of Stat5a/b. Although the exact role of phosphorylation at Tyr90 and Tyr114 is not well understood, these tyrosine residues do reside within the tetramerization domain which is important for protein-protein interactions and is involved in altering local chromatin structures upon DNA binding.[68] Stat3 also had an N-terminal (Tyr45) phosphosite hyperphosphorylated in p210$^{BCR\text{-}ABL}$ cells similar to the pattern observed in Stat5. These data support that the Stat proteins are differentially engaged by the BCR-ABL variants, and we identify the differential and specific residues through which these proteins are inducing this pathway.

3.10 Correlation of global phosphorylation signatures between p190$^{BCR\text{-}ABL}$/p210$^{BCR\text{-}ABL}$ in the MPP and Ba/F3 cells:

Next, we chose to interrogate phosphorylation signatures in another cellular background as a means of biological validation. Just as we did with our interactome analysis, we leveraged BCR-ABL variant expressing MPP cells. Experimental workflow for phosphorylation analysis in MPP cells is shown in Supplemental Figure 2b. The correlation of p190$^{BCR\text{-}ABL}$/p210$^{BCR\text{-}ABL}$ phosphorylation changes in Ba/F3 and MPP cells is plotted in Figure 3e and all MPP phosphorylation data are listed in Supplemental Table 4. Molecules with sites that showed the same trend of phosphorylation in both cell types and were also enriched among p210$^{BCR\text{-}ABL}$ interactors were Ubash3b (Tyr8), SHIP1 (Tyr918/Y886), Crk (Tyr108), and Shc1 (Tyr423). We also observed the hyperphosphorylation of the ubiquitin ligase activating site Tyr369/363 (discussed above)

on Cbl/Cblb in MPP p190$^{BCR\text{-}ABL}$ cells. As for phosphorylation signatures of proteins enriched for p190$^{BCR\text{-}ABL}$ interaction, we observed similar hyperphosphorylation of WASP, N-WASP, Wipf2, Itsn2, Dok1, and Nck1 in MPP as Ba/F3 cells further implicating p190$^{BCR\text{-}ABL}$ specificity for these molecules.

In addition to directly phosphorylating interacting proteins, BCR-ABL induces downstream signaling pathways as well. Many downstream tyrosine kinases were observed in both Ba/F3 and MPP cells to be differentially phosphorylated between p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$, which suggests that the two variants could be collaborating with other kinases to induce unique signaling programs. Sites on kinases such as tyrosine protein kinase Lyn (Lyn), tyrosine protein kinase Csk (Csk), tyrosine protein kinase Tec (Tec), tyrosine protein kinase Syk (Syk), and tyrosine kinase non-receptor 2 (Tnk2) were observed to be differentially phosphorylated in p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ cells (Figure 3.3e). Tyr265, located near the N-terminal side of the kinase domain in Lyn, was hyperphosphorylated in p210$^{BCR\text{-}ABL}$ cells in both cellular backgrounds and Tyr193/194 located within in the SH2 domain was conversely hyperphosphorylated in p190$^{BCR\text{-}ABL}$ cells. The p190$^{BCR\text{-}ABL}$ hyperphosphorylation at residues Tyr64 and Tyr18 of Csk, a regulator of Src family kinases, could indicate increased activation.

3.11 The integration of interactome and phosphorylation data highlights differently regulated molecules by p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$

To better characterize the findings from both analyses, we integrated the differential interactome and differential phosphoproteome data (Figure 3.4). Common protein interactors represent molecules that display similar association with p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$, as determined by BioID and SILAC-based mass spectrometry analysis.

These interactions likely represent general BCR-ABL interactions, since many of these same proteins are hyperphosphorylated in both BCR-ABL variant expressing cells, as compared to cells not expressing any BCR-ABL protein (Supplemental Table 2). Strikingly, overlaying the differential interactome and phosphoproteome data reveals a general agreement between increased interaction and increased phosphorylation demonstrating the power of this approach (Figure 3.4). For instance, in $p210^{BCR\text{-}ABL}$ cells, SHIP2, SHIP1, Ubash3b, Pik3ap1 and Crk have concurrent enrichment for interaction and phosphorylation. Similarly, in $p190^{BCR\text{-}ABL}$ cells, WASP, Wipf1, Nck1, Dok1, Itsn2 and DAB2 interacting protein (Dab2ip) have concurrent enrichment for interaction and phosphorylation. The large majority of the interactors that are not enriched for either variant have near equal phosphorylation signatures. An interesting group of molecules are those that appear to be equal interactors yet have differential $p190^{BCR\text{-}ABL}$/$p210^{BCR\text{-}ABL}$ phosphorylation signatures, such as the kinase leucine rich repeat kinase 1 (Lrrk1) and the focal adhesion localized protein Zyxin (Zyx). Another interesting group of molecules are those that show modest interaction preference to one variant and have increased phosphorylation in the same direction. Examples include the Arap1 (a GEF protein discussed above), the non-receptor tyrosine kinase Peak1, and cytokine dependent hematopoietic cell linker (Clnk) which is involved in the regulation of the B cell antigen receptor (Figure 3.4). Interestingly, in the MPP BioID experiment Clnk was found to be 2-fold enriched in $p190^{BCR\text{-}ABL}$ interaction which highlights the utility of combining both interaction and phosphorylation data to identify differentially regulated proteins.

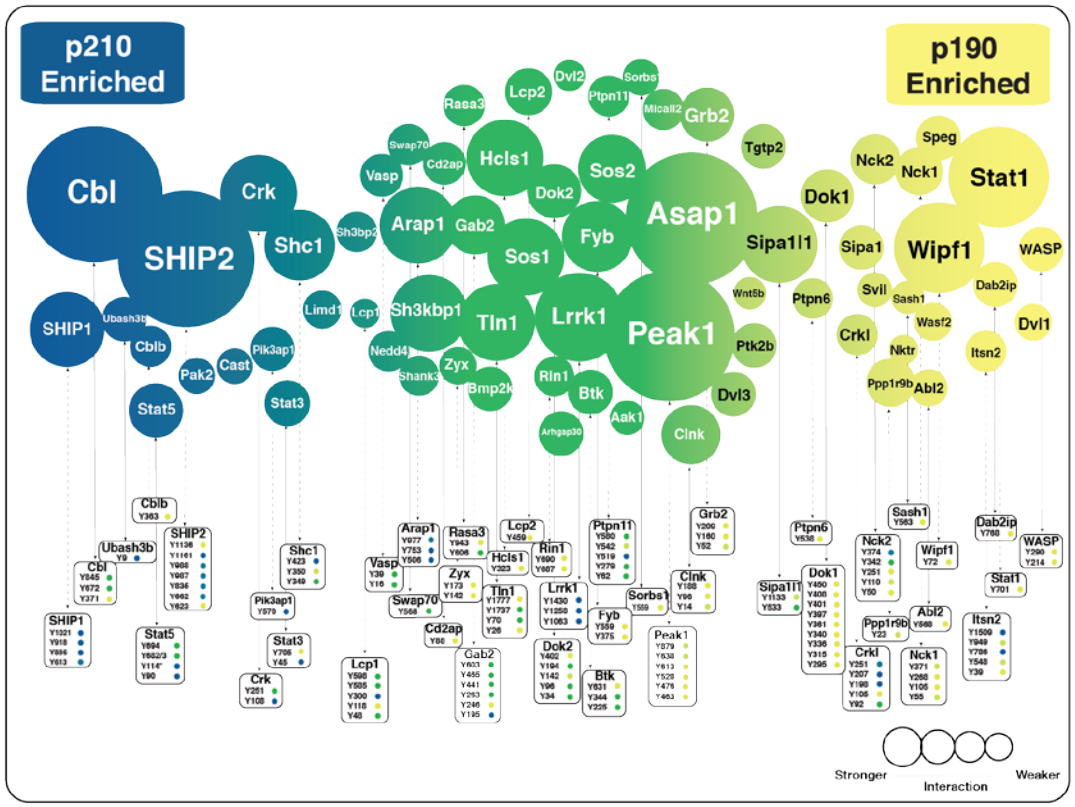

Figure 3.4 Integrative analysis of p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ phosphorylation and interactome signatures. Interacting proteins of p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ are represented by circles and are place from left to right based on their SILAC ratios in the interactome analysis. A three-color gradient represents the degree of enrichment, with blue representing an enrichment with p210$^{BCR\text{-}ABL}$ interactions, yellow representing an enrichment with p190$^{BCR\text{-}ABL}$, and green representing equal interactions between the two variants. Circles are scaled based on the number of peptide spectral matches (PSMs) in the interactome experiment, which is suggestive of the relative strength of the interactions. Arrows point toward boxes containing tyrosine sites corresponding to proteins found in interaction analysis and the phosphorylation signature observed, with respect to p190$^{BCR\text{-}ABL}$ /p210$^{BCR\text{-}ABL}$. Blue dots are used to denote at least greater than 2-fold increase in phosphorylation enrichment in p210$^{BCR\text{-}ABL}$ cells while yellow dots represents at least greater than 2-fold increase in phosphorylation in p190$^{BCR\text{-}ABL}$ cells. Lighter shades of blue and yellow dots indicate greater than 1.5-fold enriched in each cell type. Green dots signify no phosphorylation change between p190$^{BCR\text{-}ABL}$ or p210$^{BCR\text{-}ABL}$ cells.

3.12 In a multipotent background $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ regulate total protein levels of important hematopoietic drivers differently

Differential signaling programs employed by either $p190^{BCR\text{-}ABL}$ or $p210^{BCR\text{-}ABL}$ likely lead to differences in the regulation of the total proteome. In an effort to identify proteins whose abundance was differentially regulated as a result of $p190^{BCR\text{-}ABL}$ or $p210^{BCR\text{-}ABL}$ expression, we examined the total proteome data from Ba/F3 and MPP experiments (data is listed in Supplemental Tables 5 and 6, respectively). Notably, in the MPP experiment, we observed a number of important hematopoietic development/leukemia related proteins that were differentially expressed between $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$. In $p190^{BCR\text{-}ABL}$ MPP cells, many hematopoietic developmentally regulated cell surface markers were upregulated in comparison to $p210^{BCR\text{-}ABL}$ cells. These included CD14, CD81, CD9, CD166 and CD44. An important driver of B cell lymphopoiesis, transcription factor E2-alpha (E2A/Tcf3), was upregulated in $p190^{BCR\text{-}ABL}$ MPP cells. E2A is essential for initiating differentiation of uncommitted hematopoietic progenitors to pro-B cells and when translocated with the PBX1 gene is a potent driver of a subset of ALL.[69] Notable proteins upregulated in $p210^{BCR\text{-}ABL}$ MPP cells are listed in Table 2, which includes the transcription factors AT-rich interactive domain-containing protein 3A (Arid3a) and runt related transcription factors 1 and 3 (Runx1 and Runx3, respectively). Aberrant Arid3a expression levels have been shown to alter normal hematopoiesis, as overexpression of Arid3a results in inhibition of maturation of myeloid lineages while downregulation of Arid3a results in decreased B cell production.[70] Both Runx1 and Runx3 have important roles in hematopoiesis and myeloid leukemia.[71,72] Myeloperoxidase (MPO), a common

diagnostic marker differentiating myeloid from lymphoid leukemia was enriched in $p210^{BCR\text{-}ABL}$ MPP cells. Cell surface markers CD48, CD82 and the hematopoietic stem cell marker CD34 were found to be enriched in $p210^{BCR\text{-}ABL}$ MPP cells.

CHAPTER 4.

BioSITe: A METHOD FOR DIRECT DETECTION AND QUANTITIATION OF SITE -SPECIFIC BIOTINYLATION

The following work was published in the *Journal of Proteome Research* and all reference to supplemental data/figures can be found at the *Journal of Proteome Research* website with the following this citation[73].

4.1 Proximity dependent labeling technologies do not detect biotinylated proteins

To improve the biotinylated peptide detection issue, we have developed a strategy designated **Bio**tinylation **S**ite **I**dentification **Te**chnology (**BioSITe**), which is based on the use of anti-biotin antibodies to directly capture and identify biotinylated peptides in a single LC-MS/MS run. We show that detection of biotinylated peptides greatly increases the confidence of candidate protein identification in proximity-dependent biotinylation methods. By providing the site of biotinylation on proteins labeled in these methods, this approach also offers a new level of information about the structural aspects of PPIs and protein topology. Finally, we describe a simple approach for quantitative BioSITe experiments through the use of isotopically labeled biotin obviating the use of metabolic labeling based quantitative strategies such as SILAC[33].

4.2 Detection of biotinylated peptides using BioSITe

Proteins that are biotinylated by proximity-dependent biotinylation methods (Figure 1A) are generally captured by streptavidin-conjugated beads followed by on-bead digestion (hereafter referred to as the conventional method) (Figure 4.1B top). We reasoned that the strong affinity between biotin and streptavidin (K_d= ~10^{-15}) [74] is a critical factor that limits elution and subsequent detection of biotinylated peptides. To overcome this, we decided to use a capture reagent with a lower affinity, i.e. a biotin-

specific antibody, which would allow for dissociation of biotin and identification of biotinylated sites (Figure 4.1B bottom). To evaluate the potential of BioSITe in improving analysis of proximity labeling interactome studies, we leveraged our previously characterized Ba/F3 cell system that was engineered to stably express the BCR-ABL oncogene (p190 variant) cloned in-frame with BirA* [75]. Based on our testing of a panel of commercially available antibodies against biotin, we chose two polyclonal antibodies (from Abcam and Bethyl Laboratories) that yielded high signal intensities from immunoprecipitation of biotinylated proteins followed by Western blotting of whole cell lysates of BirA*-p190 (Supplementary Figure 1A). Tryptic peptides generated from these samples were separately immunoprecipitated with each antibody and analyzed by LC-MS/MS (see Methods and Materials for details regarding the enrichment and MS analysis). Peptides containing biotinylated lysines were readily detected including a biotinylated peptide of SHC1 adapter protein 1 (SHC1), a direct binding partner of BCR-ABL, and its spectrum is shown in Figure 1C (9, 10). Lysine biotinylation was evidenced by MS/MS spectrum with three signature ion peaks as well as a series of fragmentation ions harboring biotin. The signature ions, m/z 227, m/z 310 and m/z 314, correspond to fragmented biotin,

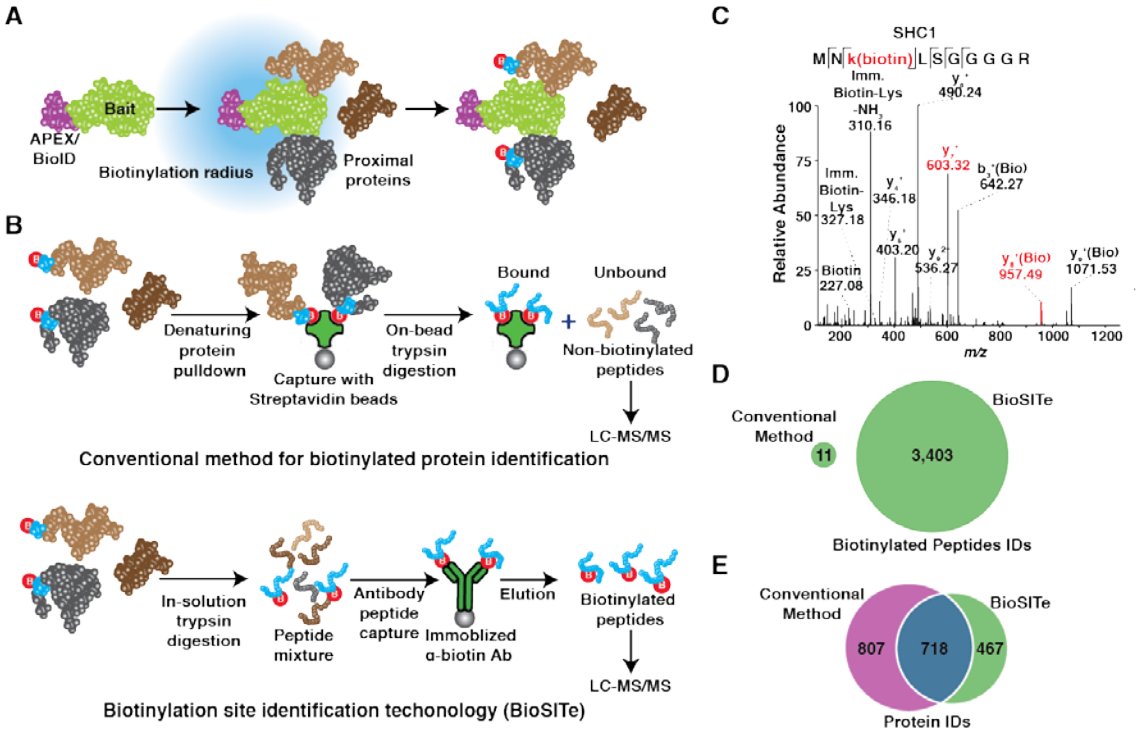

Figure 4.1 An overview of BioSITe. (A) A schematic of biotinylation labeling techniques commonly used to identify protein-protein interactions *in vivo*. A bait protein (green) fused to an engineered biotinylation enzyme (APEX2 or BioID, purple) biotinylates proteins in the proximity of the enzymes. Biotinylation - red circles with "B", interacting proteins, brown/black, labeling radius, blue. (B) In conventional methods for candidate protein identification, biotinylated proteins are generally captured by streptavidin (green) conjugated to beads under denaturing conditions. Proteins bound to the beads are subsequently digested generating non-biotinylated peptides (brown and grey), which readily elute from the beads and can be identified by LC-MS/MS, and biotinylated peptides (cyan) that remain tightly bound to the beads. In BioSITe, proteins are digested prior to enrichment and biotinylated peptides are captured using by anti-biotin antibodies coupled to beads. (C) MS/MS spectra of SHC-transforming protein 1 (SHC1) biotinylated peptide detected by LC-MS/MS. Fragment ions adjacent to the biotin modification that confirm the site of biotinylation are indicated in red. (D) Lack of overlap of biotinylated peptides identified from BirA*-BCR-ABL detected by the conventional method or BioSITe. (E) Overlap of biotinylated proteins identified by BioSITe (green) and by the conventional method of on-bead digestion (purple).

an immonium ion harboring biotin with a loss of NH3 and an immonium ion harboring biotin, respectively (4, 5). Since we observed a partially overlapping pattern in terms of site identification, we used both antibodies together in subsequent BioSITe experiments (Supplemental Table 1) (Supplementary Figure 1B).

We next sought to compare the performance of BioSITe with other methods available for biotinylated protein or peptide capture. We analyzed triplicate samples from (i) whole protein capture followed by on-bead digestion; (ii) biotinylated peptides eluted from beads subsequent to on-bead digestion; (iii) neutravidin-based biotinylated peptide capture method (DiDBiT); and, (iv) whole protein capture with anti-biotin antibodies followed by elution and in-solution digestion. MS data from BioSITe experiments identified 3,403 biotinylated peptides corresponding to 1,193 proteins (Supplemental Table 2). On-bead digestion identified 11 biotinylated peptides (Supplemental Table 3), with no overlap of those biotinylated peptide identified by BioSITe (Figure 4.1D). Similarly, elution after on-bead digestion identified 55 biotinylated peptides (Supplementary Figure 1C and Supplemental Table 4), DiDBiT detected 924 biotinylated peptides (Supplementary Figure 1D and Supplemental Table 5) and whole protein capture with anti-biotin antibodies followed by elution and in-solution digestion identified only 44 biotinylated peptides (Date not shown). Importantly, BioSITe led to identification of the largest number of biotinylated peptides of all methods tested in our study. Based on these findings, we conclude that ease of dissociating biotinylated peptides from anti-biotin antibodies combined with the peptide level enrichment is responsible for a higher number of identification using BioSITe. However, given that DiDBiT and BioSITe outperform both whole protein enrichment methods (21-85 and 77-

309 fold increase in biotinylated peptide identification, respectively), we conclude that the peptide level enrichment step provides the greatest increase in identification while the choice of capture reagent (avidin analog or antibody) provides an additional ~4-fold increase in the number of biotinylated peptides that are identified.

When we compared the list of proteins whose biotinylated sites were detected by BioSITe with proteins identified from conventional on-bead digestion, we found a substantial overlap (718 proteins) which included well-documented interacting proteins such as SHC1, ABL interactor 1 (ABI1), ubiquitin-associated and SH3 domain-containing protein B (UBASH3B), phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1 (SHIP1), inositol polyphosphate phosphatase like 1 (SHIP2) and growth factor receptor-bound protein 2 (GRB2) (Figure 4.1E) [75,76]. There were 467 biotinylated proteins identified only by BioSITe, representing a subset of biotinylated proteins that were missed by conventional on-bead digestion (Figure 4.1E). Importantly, 23% (110/467) of these proteins have more than 2 sites of biotinylation; however, the majority, 77% (359/467), of these proteins are represented by only 1 site of biotinylation indicating that these could represent low abundant proteins or transient interactions only observable by BioSITe. In addition, we observed that 807 proteins were uniquely identified by the conventional method and we hypothesize that many of these represent non-specifically bound proteins (46% (378/807) are represented by $\leq$2 PSMs) (Supplemental Table 2).

4.3 Mapping of biotinylation sites in proteins

To study the pattern of protein biotinylation and to determine if any interesting features could be revealed by the degree of biotinylation and the location of biotinylation

sites, we generated a curated list of core interactors of BCR-ABL. We then grouped these proteins according to the proportion of lysines that were detected to be biotinylated (Figure 4.2A). Among the proteins with the greatest degree of biotinylation, we found that established direct BCR-ABL interactors such as GRB2 [77,78] , GRB2 associated binding protein (GAB2) [79,80] and SHC1 [81] contained >23% of lysines that were biotinylated. The bait protein, BCR-ABL, exhibited even higher biotinylation as 75 of 95 lysine residues were biotin-modified, whereas other proteins known to interact with BCR-ABL via an adaptor subunit, e.g. PIK3CA/B, exhibited a lower extent of biotinylation (3%). Thus, we hypothesize that a high degree of biotinylation could reflect proximity and high residence time between BCR-ABL and its interactors and/or substrates. Mapping of biotinylated lysines onto the structures of corresponding proteins revealed that the majority of biotinylated lysines were in regions with low secondary structures (i.e. located within N- or C-termini or interdomain linkers outside of folded domains). As expected in this case, a majority of biotinylated sites were positioned near known functional domains involved in protein-protein interactions with BCR-ABL such as src homology 2 (SH2) domain, phosphotyrosine binding (PTB) domain and src homology 3 (SH3) domain (Figure 4.2B). For instance, GRB2, which binds BCR-ABL via its SH2 domain is highly biotinylated in both of its SH3 domains that are located adjacent to the SH2 domain, while the adapter molecule, CRK proto-oncogene adaptor protein (CRK), is additionally biotinylated on the SH2 domain itself (Figure 4.2B). Two biotinylated lysine residues (K689, K696) of signal transducer and activator of transcription 5A (STAT5A), a prominent substrate of BCR-ABL, are also found adjacent to the SH2 domain (Figure 4.2B), which mediates dimerization subsequent to

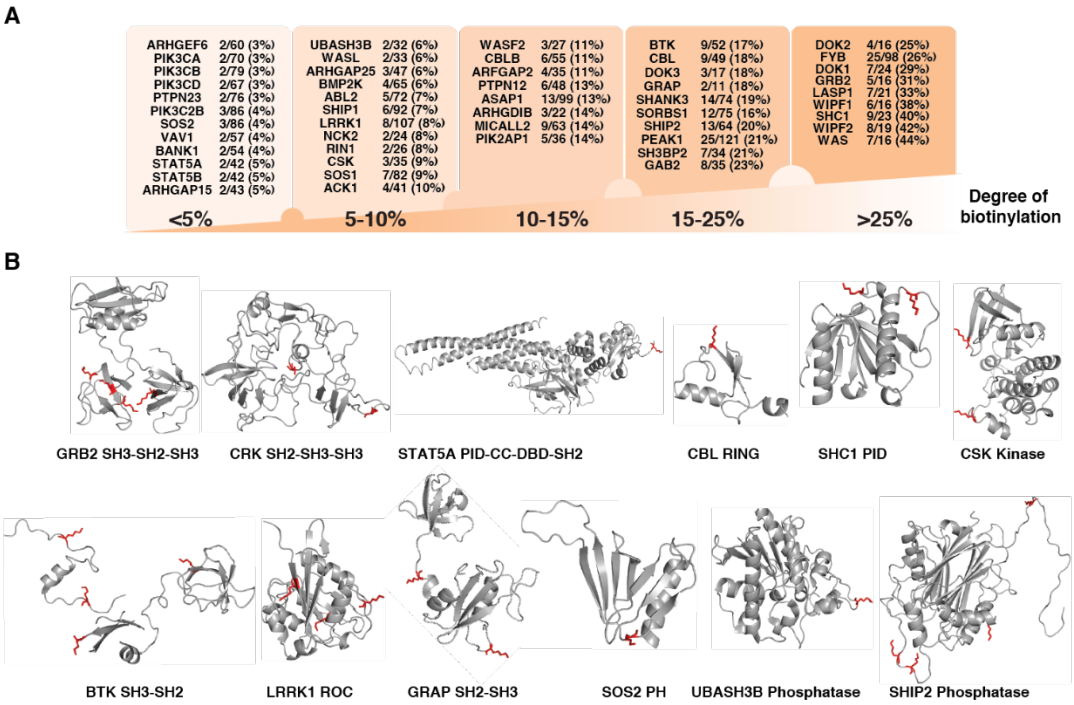

Figure 4.2 Mapping of biotinylation sites. (A) Biotinylated proteins identified by BioSITe were grouped by the degree of biotinylation. (B) Three dimensional models of representative proteins identified in the study. GRB2, and CRK are homology models based on their human homologues (PDB ID 1GRI and 2EYZ, respectively) while the structure of STAT5A was taken from PDB ID 1Y1U. The other domain structures from each protein were modeled using their human homologs from PDB. Lysine residues that are biotinylated upon interaction with BirA*-BCR-ABL are highlighted as red sticks and labeled with their position. The biotinylated lysines are in colored in red and the functional domains are indicated (CC: coiled coil, ROC: ras of complex proteins PH: pleckstrin homology, SH3: Src-homology 3, SH2: Src-homology 2, PID: phosphotyrosine interaction domain, DBD: DNA-binding domain).

phosphorylation of STAT5A by BCR-ABL [82,83]. Taken together, our observations suggest that the site level information generated by BioSITe could be potentially useful for gaining spatial information on the architecture of protein complexes.

4.4 Application of BioSITe for quantitative proteomics

Given the utility of identifying site-specific biotinylation sites, we wished to adopt a quantitative approach that would allow us to investigate relative abundance of biotinylated peptides under different biological conditions. Traditionally, accurate relative quantitation approaches have relied upon SILAC which is based on metabolic incorporation of stable isotope containing amino acids into proteins in living cells [33]. Here, instead of using heavy amino acids, we decided to use heavy biotin in the BioID experiment that would serve the same purpose. Again, we used our previously established system for interrogating the differential interactomes of two variants of BCR-ABL oncogene [75]. The p190 and p210 versions of BCR-ABL differ from each other only by the presence or absence of DH and PH domains but have distinct phenotypes in humans. We previously tagged both variants with BirA* (hereafter referred to as BirA*-p190 and BirA*-p210) [75] and used SILAC-based quantitation to characterize their differential interactomes [75,76]. We repeated the experiment by culturing the cells overnight in "heavy ($^2H_4$)" or "light (H_4)" biotin (Figure 4.3A). LC-MS/MS analysis of mixed cell lysates from BirA*-p190 and BirA*-210 cells using BioSITe revealed the expected mass differences at the MS as well as MS/MS levels on biotinylated peptides (Supplementary Figure 2A). As expected, the biotinylation sites from the DH and PH domain of BCR-ABL p210 were highly enriched in BirA*-p210 harboring cells (Figure 4.3B, Supplemental Table 6). To compare this BioSITe experiment with labeled biotin with our

previous SILAC-based BioID study, we averaged all biotinylation site ratios in the BioSITe data for specific proteins. We found many differential interactors determined by our previous studies to be relevant for BCR-ABL p190 and p210 specific signaling pathways - these included the p210 enriched interactors of SHIP1, SHIP2, CBL, UBASH3b (Supplementary Figure 4.2B) and the p190 enriched interactors of Wiskott-Aldrich Syndrome protein (WAS), and WAS/WASL-interacting protein family member 1 (Wipf1) [75,76]. Our quantitative BioSITe analysis also led to identification of potentially novel interactors that preferentially interacted either with p190 (e.g. SRSF protein kinase 2 (SRPK2) and BRCA1 associated protein (BRAP)) or with p210 (e.g. C-terminal src kinase (CSK) and PDGFA associated protein 1 (PDAP1) (Figure 4.3B). Interestingly, we did not previously identify CSK as a p210-specific interactor although it was found to be hyperphosphorylated in cells expressing p210. Thus, the preferential interaction between CSK and p210 BCR-ABL, which observed by quantitative BioSITe, further implicate its role in p210 specific signaling pathways [75,76].

Our quantitative BioSITe approach has potential advantages over traditional SILAC for proximity based labelling methods. While SILAC requires multiple passages to ensure complete metabolic labeling, which has its limitations with difficult to culture and slowly dividing cells, the addition of heavy labeled biotin can be completed with an overnight incubation in the setting of BioID. Using heavy biotin-phenol in the APEX system coupled with BioSITe is another potential of the method as APEX requires a very short labeling time (<1 minute) [3,84].

4.5 Applying BioSITe to the APEX system

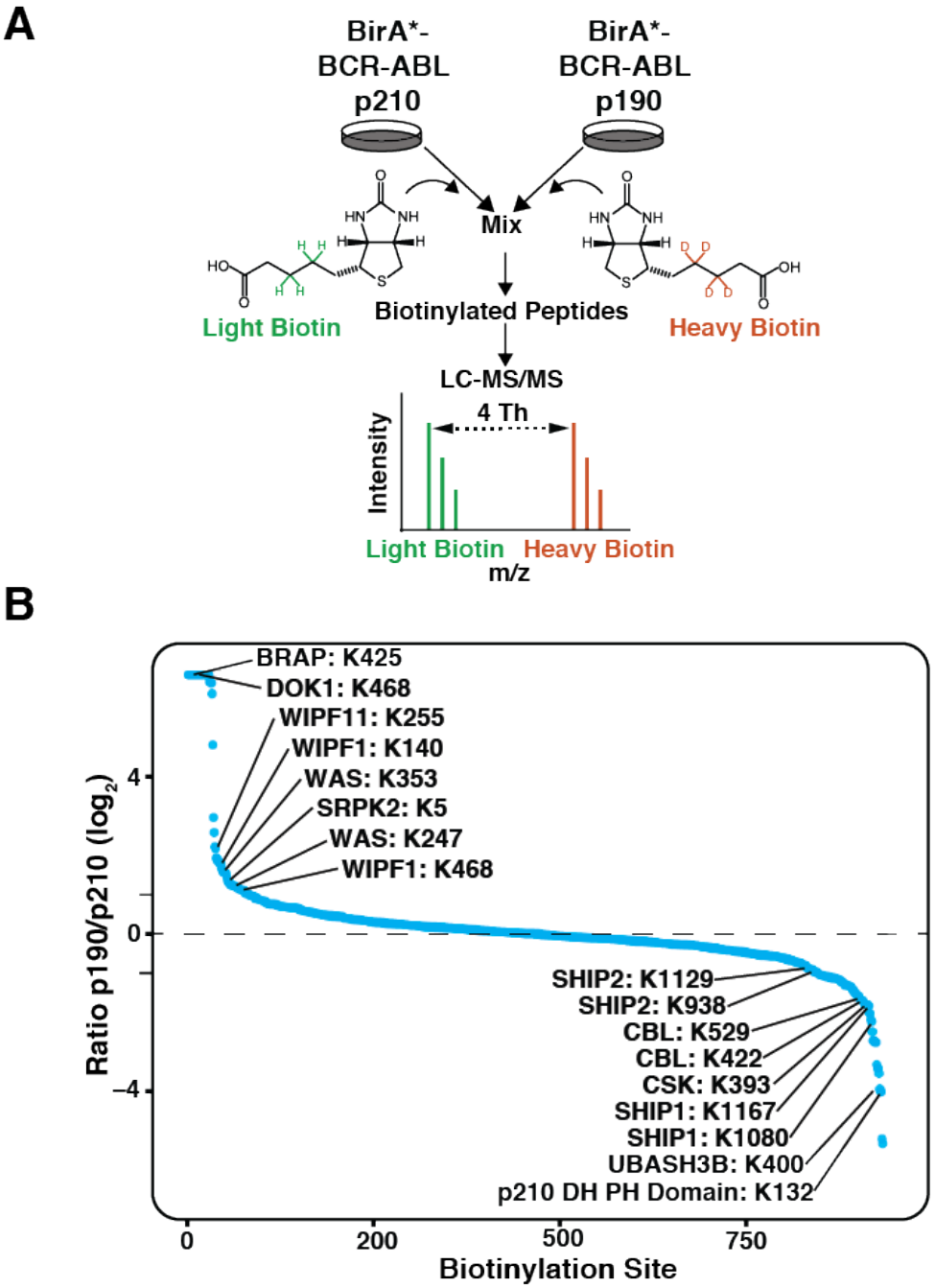

Figure 4.3 Overview of quantitative BioSITe. (A) Experimental workflow for differential interactome analysis of BirA*-p210 and BirA*-p190. Ba/F3 cells expressing BirA*-p210 and BirA*-p190 were incubated overnight with media containing light (green) or heavy biotin containing four deuterium atoms (red), respectively. Equal amounts of cell lysates from each condition were mixed and digested into peptides. Biotinylated peptides were enriched using BioSITe and analyzed by LC-MS/MS. (B) A plot of relative abundance of the biotinylated sites between BirA*-p210 and BirA*-p190. Identified biotinylated sites and the corresponding proteins are plotted according to their $\log_2$ intensity ratios (BirA*-p190/BirA*-p210).

To demonstrate the broad applicability of BioSITe, we applied our method to APEX, another proximity-dependent biotinylation method [4]. We chose an established system of mapping the mitochondrial subcellular proteome in which we expressed APEX2-fused to signal sequences targeting it to either the intermembrane space in the mitochondria (IMS -APEX2) or the cytoplasm (NES-APEX2) in HEK293T cells (Figure 4.4A)[4]. We carried out BioSITe experiments in triplicate for both IMS-APEX2 and NES-APEX2 and mass spectrometry analysis confirmed enrichment of biotin-phenol modified peptides (Figure 4.4B). Peptides containing biotin-phenol modified tyrosines were readily detected by LC-MS/MS analysis in both experiments. Biotin-phenol modified tyrosines were confirmed by MS/MS spectrum with three signature ion peaks as well as a series of fragmentation ions harboring biotin. The signature ions, 227.08, 480.19 and 497.22, correspond to fragmented biotin, biotin-phenol modified tyrosine immonium ion with loss of NH_3, and biotin-phenol modified tyrosine immonium ion, respectively (Figure 4B) [3,84].

BioSITe identified a total of 1,454 biotin-phenol modified peptides derived from 656 proteins from IMS-APEX2 labeled lysates (Supplemental Table 7) and 1,924 biotin-phenol modified peptides derived from 786 proteins from NES-APEX2 labeled lysates (Supplemental Table 8). After filtering out modified peptides common to both NES-APEX2 and IMS-APEX2 samples, we identified 384 BPMPs representing 244 proteins uniquely identified by IMS-APEX2 (Figure 4.4C) When we compared our data with a previously published study using the same system in which 135 IMS proteins were identified [4], we observed that 62 proteins were identified by both methods. Of the 182 proteins that were only identified by the BioSITe method, 74 proteins are already

annotated in MitoCarta, a mitochondrial protein database (Figure 4.4C) [85] suggesting that these proteins were likely *bona fide* proteins that were missed by the previous study. Further, BioSITe identified 345 biotinylated sites while the previous study found 80 sites of which 45 biotinylated peptides were identified by both methods and the large majority were unique to the BioSITe method (Figure 4.4D).

Recently, a method based on desthiobiotin-phenol, a synthetic biotin-phenol derivative, has been applied to the APEX system, which takes advantage of weaker affinity of desthiobiotin-phenol to streptavidin [84]. In this study, desthiobiotin-phenol-containing peptides were enriched and used to define the IMS proteome. When we compared proteins and peptides identified by all three methods (in-gel digestion-based methods by Hung et al., desthiobiotin-phenol-based study by Lee et al. and BioSITe), BioSITe yielded more labeled peptides and proteins (Supplementary Figure 3A and 3B). We have noticed that there are many proteins and peptides that are uniquely identified by each method suggesting perhaps that mitochondrial proteome is still insufficiently cataloged, issues with sampling, inherent biases of each method or a combination of these factors (Supplementary Figure 3A and 3B). As an additional caveat, we cannot rule out the possibility that each dataset includes false positives owing to porous mitochondrial outer membrane, which could itself lead to labeling of cytosolic proteins [4].

4.6 Biotinylation site-dependent topology prediction

A promising attribute of site-specific biotinylation data in the context of subcellular proteome mapping is the ability to infer topology of proteins with transmembrane domains. For example, using a construct that localizes APEX2 to the

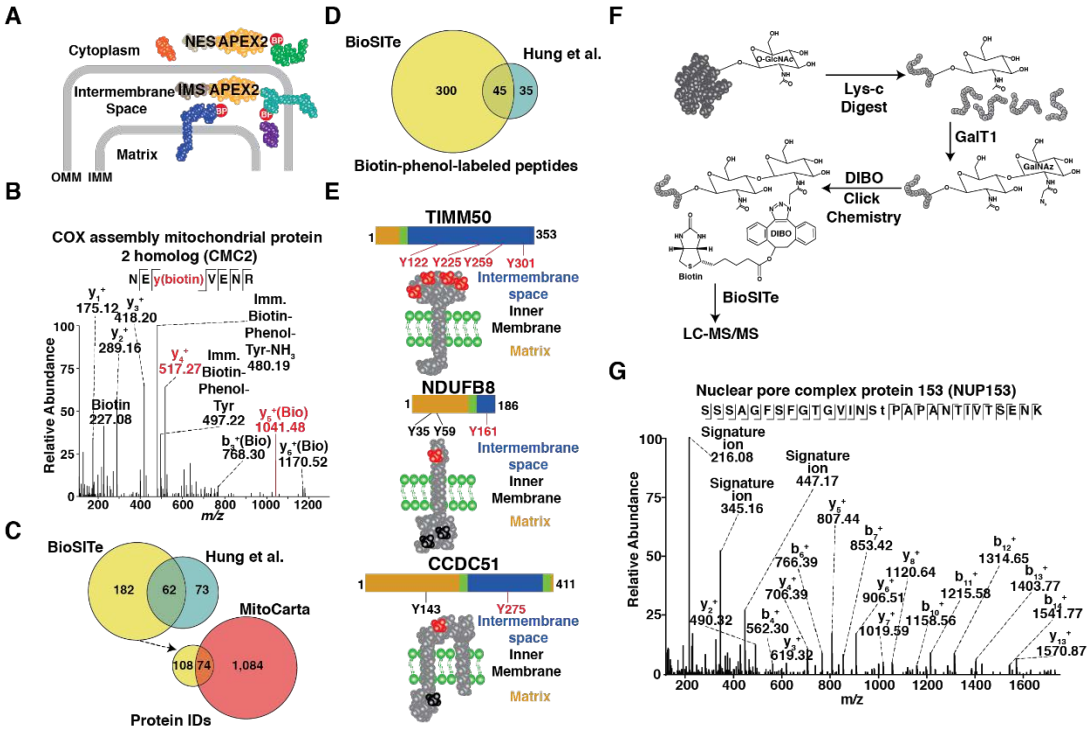

Figure 4.4 Application of BioSITe to the APEX system and biotin-based click chemistry (A) APEX2 constructs targeted to either the mitochondrial intermembrane space (IMS-APEX2) or cytoplasm (NES-APEX2) were expressed in HEK293T cells. APEX2 leads to labeling of proximal proteins in the intermembrane space or the cytoplasm with biotin-phenol (B, red). (OMM; outer mitochondrial membrane, IMM: inner mitochondrial membrane). (B) MS/MS spectrum of a biotin-phenol labeled peptide corresponding to COX assembly mitochondrial protein 2 homolog (CMC2). Peaks labeled in red indicate fragment ions series with or without mass shift by biotin modification. All peaks containing biotin are label with "Bio." (C) Overlap of proteins identified by BioSITe (yellow) with that from a previous study identifying IMS proteins using a conventional in-gel digestion method (Hung et al. blue). The proteins unique to the BioSITe IMS-APEX2 experiment were compared to published mitochondrial protein database MitoCarta 2.0 (red). (D) Overlap of biotin-phenol labeled peptides enriched by BioSITe (yellow) with that from Hung et al. (blue). (E) Predicted topology of mitochondrial transmembrane proteins of translocase of inner mitochondrial membrane 50 (TIMM50) (top), NADH:ubiquinone oxidoreductase subunit A8 (NDUFB8) (middle), and coiled-coil domain containing 51 (CCDC51) (bottom). Matrix facing (orange), transmembrane (green), and intermembrane space domains are indicated. Cartoon depictions of these proteins embedded in the inner membrane (green) are shown in grey with biotinylation sites detected by BioSITe are colored in red and black detected by Lee et al. using the Matrix-APEX2 construct. (F) Schematic overview of enrichment method for O-GlcNAcylated peptides. After LysC digestion, an azide-modified monosaccharide (GalNAz) is added to the O-GlcNAc motif using galactosyltransferase1 GalT1. Biotinylation of O-GlacNAc motif subsequently mediated by click chemistry reaction between the azide group and of GalNAz and the alkyne group of biotin DIBO alkene. Biotin is used to enrich peptides with O-GlcNAc motif using BioSITe (G) MS/MS spectra of a O-GlcNAc modified peptide corresponding to nuclear pore complex protein NUP153 (NUP153). Ion peaks at m/z 216.08, 345.15 and 447.16, labeled signature ions, correspond to fragments from click chemistry reagents.

mitochondrial matrix (Matrix-APEX2) in conjunction with the IMS-APEX2 construct, one could map the parts of transmembrane proteins that have domains in the matrix and the IMS. In some cases, BioSITe was able to provide biotinylation sites that definitively allowed us to predict the topology of 29 mitochondrial proteins. For instance, translocase of inner mitochondrial membrane 50 (TIMM50) anchored in inner mitochondria membrane, BioSITe provided biotinylation sites of Y122, Y255, Y259, and Y301 at its C-terminus (Figure 4.4E, top). In agreement with our data, the C-terminus of TIMM50 has been previously shown to be exposed to the IMS [86,87]. Biotinylation of Y301 was also described by a different study by Lee et al using desthiobiotin-phenol as a molecular probe [84]. In the case of NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8 (NDUFB8), BioSITe identified biotinylation of Y161 at the C-terminus, thus confirming the previously resolved structure [88]. Biotinylation of the IMS side in NDUFB8 was not identified by desthiobiotin-phenol-based analysis [84]. Finally, in the case of a known mitochondrial protein, coiled coil domain containing protein 51 (CCDC51), whose topology has not yet been well characterized, BioSITe led to identification of a novel site at Y275, which is positioned between the two putative transmembrane domains thereby establishing its topology in the inner membrane of the mitochondria [89]. As labeled by IMS-APEX2, we hypothesized that the region containing Y275 is located in the IMS. We also noticed that Y143 located at the N-terminus of CCDC51 has been annotated by Matrix-APEX [84], suggesting the CCDC51 is embedded in the mitochondrial inner membrane and the N terminus of CCDC51 is located inside the mitochondria matrix. Taken together, the observed site level information indicates an IMS linker bridging the

two transmembrane domains and the C terminus is also projecting into the matrix (Figure 4.4E left).

4.7 Applying BioSITe to biotin-based click chemistry

To test the potentially broad applicability of BioSITe, we applied our method to a biotin-based click-chemistry strategy to identify the post-translational modification of serine and threonine residues by N-acetylglucosamine (O-GlcNAc) [90]. O-GlcNAc is associated with a wide variety of cellular processes including cell cycle, stress response, transcription, and nutrient sensing [91,92]. To detect O-GlcNAc modified residues, whole cell lysates from HEK293T cells were digested by the protease Lys-C and the resulting peptides were subjected to PNGase F and mutant β-1,4-galactosyltransferase (Gal-T1 (Y289L)) treatment which lead to specific labeling of O-GlcNAc modified peptides with an additional sugar containing an azide (GalNAz) moiety, which is compatible with click chemistry (Figure 4.4F). These modified peptides were subsequently biotinylated in a second reaction by click chemistry using DIBO biotin alkyne where the azide group of UDP-GalNAz recruits the alkyne group attached to biotin [93,94]. Peptides containing biotin were then captured using BioSITe for LC-MS/MS analysis.

Our LC-MS/MS analysis identified O-GlcNAc sites on many known O-GlcNAc modified proteins including host cell factor 1 (HCFC1), nuclear pore complex protein 153 (NUP153), nuclear pore complex protein 214 (NUP214) and nuclear pore complex protein 98 (NUP98) [95,96] (Supplementary Table 9). A total of 10 known O-GlcNAc sites were identified on HCF1 while 5 known sites were identified on NUP153. NUP153 is a nuclear pore protein with zinc finger domains which are critical for the recruitment of COP1 (coatomer complex protein 1) to the nuclear envelope. NUP153 is involved in the

breakdown of the nuclear envelope at the prophase of cell cycle [97]. It has been shown that the level of O-GlcNAcylation at S529 increases dramatically upon phosphorylation of NUP153 at S534 by CDK1 during cell cycle [95]. Figure 4G shows an MS/MS spectrum of an O-GlcNAc-modified peptide from NUP153. Click chemistry strategies are not limited to identification sites of glycosylated proteins but have many applications, in particular, recent studies of small molecule probes couple with click chemistry are enabling high-throughput drug discovery [98,99]. We anticipate that BioSITe could benefit these experiments as many existing reagents rely on biotin for enrichment by preclude site specific identification for the reasons described above.

CHAPTER 5.

DISCUSSION

Despite speculations regarding the roles of mammalian DGKs in modulating synaptic function, the function of DGKθ in the brain has remained unknown. The data presented in this thesis were designed to test the hypothesis that DGKθ modulates neurotransmitter release from central synapses. I found that DGKθ protein expression is coincident with synaptogenesis and localizes specifically to excitatory synapses. Both acute and chronic loss of DGKθ from cortical neurons resulted in slowed SV retrieval following neuronal stimulation. SV recycling kinetics could be rescued by ectopic expression of enzymatically active, but not inactive, DGKθ, thus, implicating DGKθ catalytic activity in promoting the efficient recycling of SVs following neuronal activity.

5.1 BCR-ABL variant signaling

Although BCR-ABL signaling has been intensively investigated for many years, a direct and systematic comparison of $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ signaling differences has never been undertaken. In this study, we directly compared $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ protein-protein interactions and global protein phosphorylation signatures in homogenous cell culture systems that model early hematopoietic differentiation. Our observations indicate many differences between $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ in signaling pathway regulation.

In 1996, Ilaria and Van Etten,[21] later confirmed in part by Druker and colleagues in 2009,[12] revealed one of the first pathways to be differentially regulated by $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$. Both groups use immunoprecipitation and anti-phosphotyrosine immunoblotting to measure the total phosphorylation status of Stat proteins induced by both variants. However, these analyses lacked tyrosine site-specific resolution, which we now provide. Our site-specific analyses provide a deeper understanding of the nature of

the differential regulation of Stat proteins. Specifically, our analysis identified $p190^{BCR\text{-}ABL}$/$p210^{BCR\text{-}ABL}$ differential regulation of sites within the trans-activating domains of Stat proteins, that when phosphorylated, lead to dimerization and subsequent DNA binding. Our interaction analysis leads us to conclude that these differential phosphorylation signatures are likely mediated by differential interactions between Stat proteins and the two variants. Stat proteins have a major role in normal hematopoiesis – for example, when Stat1 is deleted in mice, B lymphoid development is significantly reduced.[100] This is potentially relevant to the $p190^{BCR\text{-}ABL}$ interaction and phosphorylation signatures observed in our study as $p190^{BCR\text{-}ABL}$ could be engaging Stat1 differently than $p210^{BCR\text{-}ABL}$ to drive the associated B cell phenotype. Furthermore, Stat proteins have been revealed to be incredibly important for BCR-ABL mediated leukemogenesis. Specifically, both Stat3 and Stat5, found to exhibit preference for $p210^{BCR\text{-}ABL}$ interaction, are indispensable for the initiation of $p210^{BCR\text{-}ABL}$ leukemia and Stat5 is necessary for maintaining CML.[101,102] In addition, Janus kinase 2 (Jak2), the upstream kinase targeting Stat5, has also been shown to be required for $p190^{BCR\text{-}ABL}$ to initiate leukemic transformation, while Jak2 is dispensable for $p210^{BCR\text{-}ABL}$ transformation[55] suggesting that $p190^{BCR\text{-}ABL}$/$p210^{BCR\text{-}ABL}$ differential activation of Stat proteins is necessary for different phenotypes. This is particularly relevant to our findings as the preferential direct interaction of Stat5 with $p210^{BCR\text{-}ABL}$ could circumvent the need of Jak2, where $p190^{BCR\text{-}ABL}$ requires an intermediary, Jak2, for leukemia initiation because it lacks this robust Stat5 interaction. We conclude Stat protein differential interaction, and potentially activation, associated with $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$appears to play a role in how the two variants drive

leukemia, in particular the apparent lineage preference (AML versus CML), and certainly warrants further investigation.

Importantly, our analyses uncovered a potential differential subcellular localization as a feature of $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ unique signaling programs (Figure 5). Because $p190^{BCR\text{-}ABL}$ lacks the PH domain, which mediates interactions with the plasma membrane, it may be free to associate more with the cytoskeleton. In support of this, we observed increased interaction of $p190^{BCR\text{-}ABL}$ with cytoskeletal modifiers such as Wiskott-Aldrich syndrome family members, hyperphosphorylation of many cytoskeletal proteins as well as others such as Dok1, Nck1, Nck2, Grb2, and SHIP2 that are involved in cytoskeletal reorganization.[59,60] Differential interaction with cytoskeleton has previously been implicated in studies showing that $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ induced distinctly different cellular motilities.[103,104] Strikingly, in support of our findings, previous data provided evidence that deletion of the C-terminal F-actin binding domain in $p190^{BCR\text{-}ABL}$ reduced its oncogenicity, while the same deletion in $p210^{BCR\text{-}ABL}$ appeared dispensable for establishing leukemia.[46,47]

In contrast to an apparent preference of $p190^{BCR\text{-}ABL}$ with cytoskeletal components and regulators, $p210^{BCR\text{-}ABL}$, which contains the PH domain, tends to interact with molecules that suggest its localization to be more proximal to the plasma membrane (Figure 5). We identified $p210^{BCR\text{-}ABL}$ enriched interaction with Cbl, Cblb, SHIP1, SHIP2, Shc1, and Ubash3b which are all involved in the regulation of plasma membrane bound receptors.[51,105,106] Ubash3b is a particularly interesting molecule that could be playing some role in attenuating certain functions of $p210^{BCR\text{-}ABL}$, including SHIP1 degradation by inhibiting Cbl and binding to ubiquitinated proteins. Moreover, Ubash3b has been

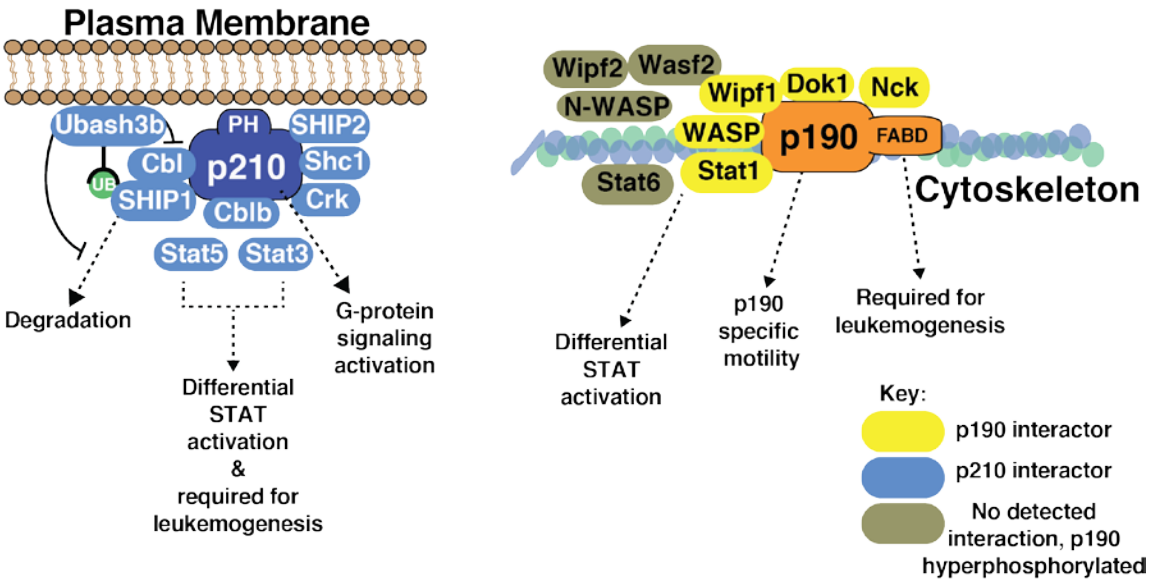

Figure 5. Model of differential p190$^{BCR\text{-}ABL}$ and p210$^{BCR\text{-}ABL}$ signaling. Interaction, determined by BioID interactome analysis, is by close proximity of ellipses. Yellow coloring indicates proteins that are enriched in p190$^{BCR\text{-}ABL}$ interaction and blue coloring denote p210$^{BCR\text{-}ABL}$ enriched interaction. Green/brown coloring indicates no detected interaction in BioID but are place near p190$^{BCR\text{-}ABL}$ based on their p190$^{BCR\text{-}ABL}$ hyperphosphorylation signature. p210$^{BCR\text{-}ABL}$ is shown in dark blue to be anchored to the plasma membrane by the PH domain. p190$^{BCR\text{-}ABL}$ is shown in orange to be anchored to the cytoskeleton by the F-acting binding domain interacting with molecules therein. Dotted lines with arrowheads refer to molecular functions observed in previous studies. Line with semicircle arrowhead represents the ubiquitin binding function of Ubash3b. Ubiquitin modifications (UB) are represented by green circles.

shown to have phosphatase activity toward some Src and Syk family tyrosine kinases[52,107] and we found evidence of dysregulation of both Lyn and Syk in our analysis. Src kinases have been shown to be required for the establishment of BCR-ABL positive B-ALL while being dispensable for CML establishment in mouse models, stressing the importance of our data describing the p210$^{BCR-ABL}$-specific preference for interactions with and regulation of Ubash3b and its downstream regulation of Src kinases.[108]

Another interesting level of differential p190$^{BCR-ABL}$/p210$^{BCR-ABL}$ regulation was found at the level of total proteome control. In the MPP experiment we found a surprising amount of developmental and leukemia related proteins differentially regulated in p190$^{BCR-ABL}$ and p210$^{BCR-ABL}$ expressing cells. Particularly, the presence of the myeloid/myeloid neoplasm centric proteins such as Runx1, Runx3, and myeloperoxidase associated with p210$^{BCR-ABL}$ cells and the presence of B cell centric proteins such as Tcf3 and Pbxip1 associated with p190$^{BCR-ABL}$ cells is interesting. The presence of CD34 on p210$^{BCR-ABL}$ and not p190$^{BCR-ABL}$ MPP cells is perhaps consistent with the observation that p190$^{BCR-ABL}$ requires a B cell differentiation event from the stem cell compartment to establish leukemia, while p210$^{BCR-ABL}$ leukemogenesis continually originates from stem cell compartment, in mouse models.[109]

Reckel et al.[110] have reported a similar study in this issue where they also examine differential signaling by p190$^{BCR-ABL}$ and p210$^{BCR-ABL}$. Both studies have identified many of the same molecules to be differentially regulated by p190$^{BCR-ABL}$ and p210$^{BCR-ABL}$, underscoring the reproducibility of global proteomic approaches to study signaling pathways in two different laboratories. However, there are also some minor differences that we believe are attributable to two broad areas - technical and biological. The

methods employed by each group had some differences; for example, different phosphopeptide enrichment strategies, mass spectrometers, search algorithms, quantification methods and methods for interactome analysis (BioID vs. affinity purification) were used. Of the two approaches used for interactome analysis, the BioID system is likely to detect direct interactions and might allow for detection of more transient interactions, while affinity purification methods might favor stronger biochemical interactions. Thus, taken together, both approaches to detect protein-protein interactions are somewhat complementary, as each method has its own biases for particular subsets of interactions, as has been previously shown.[8] An important contributing factor to biological variation observed between the two studies is likely the relative expression level of the two variants in Ba/F3 cells.

Regardless of the differences, the two studies were both able to highlight many of the same molecules and pathways, which have major relevance to understanding BCR-ABL biology. Most prominent of which are the identification of $p210^{BCR\text{-}ABL}$ specific interaction of Ubash3b/Sts-1 by both affinity purification and BioID, differential regulation of members of the STAT and Src families, and the association of $p190^{BCR\text{-}ABL}$ with cytoskeletal elements, leading to a general conclusion that the variants may exhibit different subcellular localization or enrichment. The differential phosphorylation signatures of Stat and Src family members are particularly attractive, as inhibition of these molecules could be further evaluated in the context of current tyrosine kinase inhibitor-based therapies targeting BCR-ABL, a strategy that would be especially important in the treatment of BCR-ABL positive ALL.

In summary, we have hypothesized that the disparate clinical and experimental phenotypes associated with the BCR-ABL variants likely originate from inherent differences in signaling. Our findings in Ba/F3 and MPP cell lines represent a first glimpse into the complexity of how BCR-ABL variants differentially interface with hematopoiesis and drive leukemogenesis via their unique signaling programs. We hypothesize that understanding how $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ differently regulate key signaling molecules will help explain the lineage preferences and lead to new therapeutic inroads. Our data provide a clear set of differently regulated molecules to test in specific hematopoietic stages and in more clinically directed assays.

5.2 BioSITe conclusions and future of interactomics

We have shown that BioSITe permits enrichment and detection of biotinylated peptides including the localization of biotinylation sites within peptides. Using an isotopically labeled version of biotin, BioSITe allows for a quantitative analysis that is ideal for characterizing molecular differences across different biological conditions. Although, we have demonstrated the use of a four Dalton heavier version of biotin, which is commercially available, other biotin isotopologues can be generated that will allow for higher levels of multiplexing. The ability to obtain site level information in terms of biotinylation can provide spatial information regarding the architecture of protein complexes and topology of subcellular proteomes. We have also shown the broad applicability of BioSITe by applying this strategy to a biotin-based click chemistry strategy for identifying O-GlcNAc modified sites. Overall, BioSITe is an attractive option as a simpler strategy with a high yield than previously used methods, which relied on low yields of biotinylated peptides and, in some cases, a higher number of LC-MS/MS

runs in-gel digests or customized molecular probes (e.g. desthiobiotin-phenol), which are out of the reach for most laboratories to generate. Additionally, the site level information provided in click chemistry approaches will circumvent the need for complicated cleavable tags that are frequently employed. Taken together, we foresee the general applicability of BioSITe in most applications that involve biotinylation as a strategy to tag proteins or post-translational modifications. While this paper was being finalized for publication, a report on a similar strategy using an antibody approach for direct enrichment of biotinylated peptide was published[35].

The effect BioSITe will have on the field of proximity dependent biotinylation assays still has some unknowns. The degree to which the site of biotinylation can predict the point of molecular contact of a PPI remains unclear. With the accumulation of more BioSITe datasets it is possible that trends will arise. It is clear, however, the utility of the site of biotinylation has for the combination of membrane separated subcellular proteome mapping and BioSITe for predicting unknown trans-membrane containing protein topology.

Beyond the site of biotinylation information provided by BioSITe the entirety of proximity dependent biotinylation methods in general have important implications of the future of detecting of PPIs. Co-immunoprecipitation (Co-IP) experiments have traditionally been the method of choice for detecting PPIs. However, the overlap between BioID, for example, and Co-IP experiments has been shown to be more "complimentary" than overlapping[8]. Comparing the BCR-ABL interactomes determine by Co-IP experiments by Reckel et al. and the BioID experiments described in this thesis both carried out in the same cellular system we observe the same lack of overlap (Figure 6).

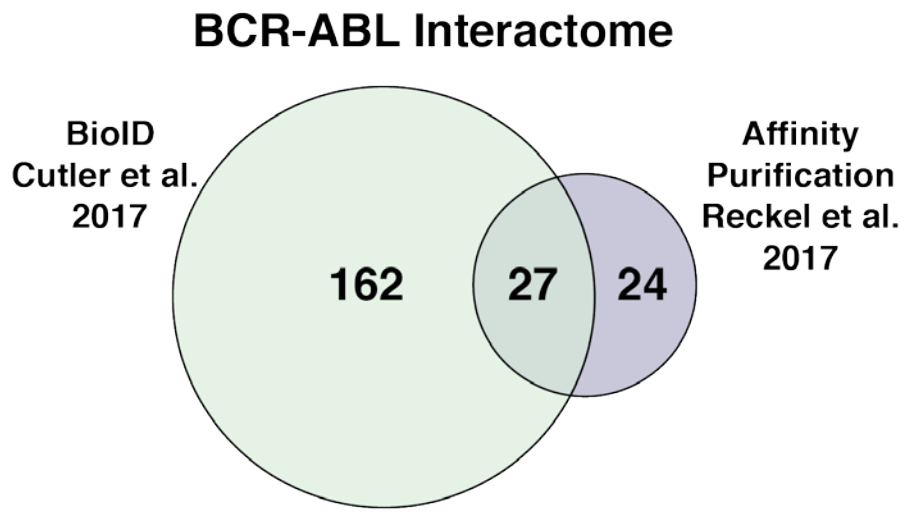

Figure 6. Venn diagram comparing Cutler et al. with Reckel et al. BioID and affinity purification interactome studies of BCR-ABL

BioID identified many more proteins than Co-IP, there are only 27 overlapping identities, and there are proteins uniquely identified in both methods, 162 for BioID and 24 for Co-IP. Moreover, only one protein, Ubash3b, was found to be differentially interacting with $p190^{BCR\text{-}ABL}$ and $p210^{BCR\text{-}ABL}$ cells in both studies[41,111]. This suggests that proximity dependent labeling technologies and Co-IP experiments measure different properties a given protein's interactome. Perhaps Co-IP measures the stronger "biochemical" interactions and the proximity-based methods capture where the protein has been and what are its neighbors. It is incumbent on the field, through downstream biological experiments, to test putative interactors identified in both Co-IP and proximity-based methods that are essential to the function of the bait protein being studied.

BIBLIOGRAPHY

1 Hunter T. Signaling—2000 and Beyond. *Cell* 2000; **100**: 113–127.

2 Kim DI, Roux KJ. Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends Cell Biol* 2016; **26**: 804–817.

3 Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA *et al.* Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* 2013; **339**: 1328–1331.

4 Hung V, Zou P, Rhee H-W, Udeshi ND, Cracan V, Svinkina T *et al.* Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging. *Mol Cell* 2014; **55**: 332–341.

5 Chakravartty V, Cronan JE. The Wing of a Winged Helix-Turn-Helix Transcription Factor Organizes the Active Site of BirA, a Bifunctional Repressor/Ligase. *J Biol Chem* 2013; **288**: 36029–36039.

6 Chapman-Smith A, Cronan JE. The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci* 1999; **24**: 359–363.

7 Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 2012; **196**: 801–810.

8 Lambert J-P, Tucholska M, Go C, Knight JDR, Gingras A-C. Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *J Proteomics* 2015; **118**: 81–94.

9 Han X, Aslanian A, Yates JR. Mass Spectrometry for Proteomics. *Curr Opin Chem Biol* 2008; **12**: 483–490.

10 Van Etten RA. Aberrant cytokine signaling in leukemia. *Oncogene* 2007; **26**: 6738–6749.

11 Score J, Calasanz MJ, Ottman O, Pane F, Yeh RF, Sobrinho-Simões MA *et al.* Analysis of genomic breakpoints in p190 and p210 BCR–ABL indicate distinct mechanisms of formation. *Leukemia* 2010; **24**: 1742–1750.

12 Demehri S, O'Hare T, Eide CA, Smith CA, Tyner JW, Druker BJ *et al.* The function of the pleckstrin homology domain in BCR–ABL-mediated leukemogenesis. *Leukemia* 2009; **24**: 226–229.

13 Foà R, Vitale A, Vignetti M, Meloni G, Guarini A, Propris MSD *et al.* Dasatinib as first-line treatment for adult patients with Philadelphia chromosome–positive acute lymphoblastic leukemia. *Blood* 2011; **118**: 6521–6528.

14 Gurion R, Raanani P, Vidal L, Leader A, Gafter-Gvili A. First line treatment with newer tyrosine kinase inhibitors in chronic myeloid leukemia associated with deep and durable molecular response – systematic review and meta-analysis. *Acta Oncol* 2016; **55**: 1077–1083.

15 Tala I, Chen R, Hu T, Fitzpatrick ER, Williams DA, Whitehead IP. Contributions of the RhoGEF activity of p210 BCR/ABL to disease progression. *Leukemia* 2013; **27**: 1080–1089.

16 Li S, Ilaria RL, Million RP, Daley GQ, Etten RAV. The P190, P210, and P230 Forms of the BCR/ABL Oncogene Induce a Similar Chronic Myeloid Leukemia–like Syndrome in Mice but Have Different Lymphoid Leukemogenic Activity. *J Exp Med* 1999; **189**: 1399–1412.

17 Lugo TG, Pendergast AM, Muller AJ, Witte ON. Tyrosine kinase activity and transformation potency of bcr-abl oncogene products. *Science* 1990; **247**: 1079–1082.

18 Hantschel O. Structure, Regulation, Signaling, and Targeting of Abl Kinases in Cancer. *Genes Cancer* 2012; **3**: 436–446.

19 Cilloni D, Saglio G. Molecular Pathways: BCR-ABL. *Clin Cancer Res* 2012; **18**: 930–937.

20 Harnois T, Constantin B, Rioux A, Grenioux E, Kitzis A, Bourmeyster N. Differential interaction and activation of Rho family GTPases by p210bcr-abl and p190bcr-abl. *Oncogene* 2003; **22**: 6445–6454.

21 Ilaria RL, Etten RAV. P210 and P190BCR/ABL Induce the Tyrosine Phosphorylation and DNA Binding Activity of Multiple Specific STAT Family Members. *J Biol Chem* 1996; **271**: 31704–31710.

22 Goss VL, Lee KA, Moritz A, Nardone J, Spek EJ, MacNeill J *et al.* A common phosphotyrosine signature for the Bcr-Abl kinase. *Blood* 2006; **107**: 4888–4897.

23 Rubbi L, Titz B, Brown L, Galvan E, Komisopoulou E, Chen SS *et al.* Global Phosphoproteomics Reveals Crosstalk Between Bcr-Abl and Negative Feedback Mechanisms Controlling Src Signaling. *Sci Signal* 2011; **4**: ra18.

24 Titz B, Low T, Komisopoulou E, Chen SS, Rubbi L, Graeber TG. The proximal signaling network of the BCR-ABL1 oncogene shows a modular organization. *Oncogene* 2010; **29**: 5895–5910.

25 Brehme M, Hantschel O, Colinge J, Kaupe I, Planyavsky M, Köcher T *et al.* Charting the molecular network of the drug target Bcr-Abl. *Proc Natl Acad Sci* 2009; **106**: 7414–7419.

26 De La Fuente EK, Dawson CA, Nelin LD, Bongard RD, McAuliffe TL, Merker MP. Biotinylation of membrane proteins accessible via the pulmonary circulation in normal and hyperoxic rats. *Am J Physiol* 1997; **272**: L461-470.

27 Rybak J-N, Scheurer SB, Neri D, Elia G. Purification of biotinylated proteins on streptavidin resin: a protocol for quantitative elution. *Proteomics* 2004; **4**: 2296–2299.

28 Tong X, Smith LM. Solid-phase method for the purification of DNA sequencing reactions. *Anal Chem* 1992; **64**: 2672–2677.

29 Morag E, Bayer EA, Wilchek M. Reversibility of biotin-binding by selective modification of tyrosine in avidin. *Biochem J* 1996; **316 (Pt 1)**: 193–199.

30 Laitinen OH, Nordlund HR, Hytönen VP, Uotila STH, Marttila AT, Savolainen J *et al.* Rational design of an active avidin monomer. *J Biol Chem* 2003; **278**: 4010–4014.

31 Taskinen B, Zauner D, Lehtonen SI, Koskinen M, Thomson C, Kähkönen N *et al.* Switchavidin: reversible biotin-avidin-biotin bridges with high affinity and specificity. *Bioconjug Chem* 2014; **25**: 2233–2243.

32 Schiapparelli LM, McClatchy DB, Liu H-H, Sharma P, Yates JR, Cline HT. Direct Detection of Biotinylated Proteins by Mass Spectrometry. *J Proteome Res* 2014; **13**: 3966–3978.

33 Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A *et al.* Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol Cell Proteomics* 2002; **1**: 376–386.

34 Roux KJ, Kim DI, Burke B. BioID: A Screen for Protein-Protein Interactions. In: *Current Protocols in*

Protein Science. John Wiley & Sons, Inc., 2011 (accessed 22 Apr2016).

35 Pongubala JMR, Northrup DL, Lancki DW, Medina KL, Treiber T, Bertolino E *et al.* Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nat Immunol* 2008; **9**: 203–215.

36 Medina KL, Pongubala JMR, Reddy KL, Lancki DW, DeKoter R, Kieslinger M *et al.* Assembling a Gene Regulatory Network for Specification of the B Cell Fate. *Dev Cell* 2004; **7**: 607–617.

37 Mitchell CJ, Kim M-S, Na CH, Pandey A. PyQuant: A Versatile Framework for Analysis of Quantitative Mass Spectrometry Data. *Mol Cell Proteomics* 2016; **15**: 2829–2838.

38 Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008; **26**: 1367–1372.

39 Mellacheruvu D, Wright Z, Couzens AL, Lambert J-P, St-Denis NA, Li T *et al.* The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods* 2013; **10**: 730–736.

40 Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J *et al.* STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2015; **43**: D447–D452.

41 Cutler JA, Tahir R, Sreenivasamurthy SK, Mitchell C, Renuse S, Nirujogi RS *et al.* Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. *Leukemia* 2017; **31**: 1513–1524.

42 Kim M-S, Zhong Y, Yachida S, Rajeshkumar NV, Abel ML, Marimuthu A *et al.* Heterogeneity of Pancreatic Cancer Metastases in a Single Patient Revealed by Quantitative Proteomics. *Mol Cell Proteomics* 2014; **13**: 2803–2811.

43 Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 2012; **196**: 801–810.

44 Daley GQ, Baltimore D. Transformation of an interleukin 3-dependent hematopoietic cell line by the chronic myelogenous leukemia-specific P210bcr/abl protein. *Proc Natl Acad Sci* 1988; **85**: 9312–9316.

45 Bhatia R, Munthe HA, Verfaillie CM. Role of abnormal integrin-cytoskeletal interactions in impaired β1 integrin function in chronic myelogenous leukemia hematopoietic progenitors. *Exp Hematol* 1999; **27**: 1384–1396.

46 Wertheim JA, Perera SA, Hammer DA, Ren R, Boettiger D, Pear WS. Localization of BCR-ABL to F-actin regulates cell adhesion but does not attenuate CML development. *Blood* 2003; **102**: 2220–2228.

47 Heisterkamp N, Voncken JW, Senadheera D, Gonzalez-Gomez I, Reichert A, Haataja L *et al.* Reduced oncogenicity of p190 Bcr/Abl F-actin–binding domain mutants. *Blood* 2000; **96**: 2226–2232.

48 Salgia R, Li JL, Ewaniuk DS, Pear W, Pisick E, Burky SA *et al.* BCR/ABL induces multiple abnormalities of cytoskeletal function. *J Clin Invest* 1997; **100**: 46–57.

49 Hantschel O, Wiesner S, Güttler T, Mackereth CD, Rix LLR, Mikes Z *et al.* Structural Basis for the Cytoskeletal Association of Bcr-Abl/c-Abl. *Mol Cell* 2005; **19**: 461–473.

50 Yi S-J, Lee H-T, Groffen J, Heisterkamp N. Bcr/Abl P190 interaction with Spa-1, a GTPase activating protein for the small GTPase Rap1. *Int J Mol Med* 2008; **22**: 453–458.

51 Kowanetz K, Crosetto N, Haglund K, Schmidt MHH, Heldin C-H, Dikic I. Suppressors of T-cell Receptor Signaling Sts-1 and Sts-2 Bind to Cbl and Inhibit Endocytosis of Receptor Tyrosine Kinases. *J Biol Chem* 2004; **279**: 32786–32795.

52 Mikhailik A, Ford B, Keller J, Chen Y, Nassar N, Carpino N. A Phosphatase Activity of Sts-1 Contributes to the Suppression of TCR Signaling. *Mol Cell* 2007; **27**: 486–497.

53 Ruschmann J, Ho V, Antignano F, Kuroda E, Lam V, Ibaraki M *et al.* Tyrosine phosphorylation of SHIP promotes its proteasomal degradation. *Exp Hematol* 2010; **38**: 392-402.e1.

54 Sattler M, Verma S, Byrne CH, Shrikhande G, Winkler T, Algate PA *et al.* BCR/ABL Directly Inhibits Expression of SHIP, an SH2-Containing Polyinositol-5-Phosphatase Involved in the Regulation of Hematopoiesis. *Mol Cell Biol* 1999; **19**: 7473–7480.

55 Hantschel O, Warsch W, Eckelhart E, Kaupe I, Grebien F, Wagner K-U *et al.* BCR-ABL uncouples canonical JAK2-STAT5 signaling in chronic myeloid leukemia. *Nat Chem Biol* 2012; **8**: 285–293.

56 Gwin K, Frank E, Bossou A, Medina KL. Hoxa9 Regulates Flt3 in Lymphohematopoietic Progenitors. *J Immunol Baltim Md 1950* 2010; **185**: 6572–6583.

57 Heydarian M. Prediction of Gene Activity in Early B Cell Development Based on an Integrative Multi-Omics Analysis. *J Proteomics Bioinform* 2014; **07**. doi:10.4172/jpb.1000302.

58 Massaad MJ, Ramesh N, Geha RS. Wiskott-Aldrich syndrome: a comprehensive review. *Ann N Y Acad Sci* 2013; **1285**: 26–43.

59 Benesch S, Lommel S, Steffen A, Stradal TEB, Scaplehorn N, Way M *et al.* Phosphatidylinositol 4,5-Biphosphate (PIP2)-induced Vesicle Movement Depends on N-WASP and Involves Nck, WIP, and Grb2. *J Biol Chem* 2002; **277**: 37771–37776.

60 Woodring PJ, Meisenhelder J, Johnson SA, Zhou G-L, Field J, Shah K *et al.* c-Abl phosphorylates Dok1 to promote filopodia during cell spreading. *J Cell Biol* 2004; **165**: 493–503.

61 Sattler M, Verma S, Pride YB, Salgia R, Rohrschneider LR, Griffin JD. SHIP1, an SH2 Domain Containing Polyinositol-5-phosphatase, Regulates Migration through Two Critical Tyrosine Residues and Forms a Novel Signaling Complex with DOK1 and CRKL. *J Biol Chem* 2001; **276**: 2451–2458.

62 Lamkin TD, Walk SF, Liu L, Damen JE, Krystal G, Ravichandran KS. Shc Interaction with Src Homology 2 Domain Containing Inositol Phosphatase (SHIP) in Vivo Requires the Shc-Phosphotyrosine Binding Domain and Two Specific Phosphotyrosines on SHIP. *J Biol Chem* 1997; **272**: 10396–10401.

63 Deneubourg L, Elong Edimo W 's, Moreau C, Vanderwinden J-M, Erneux C. Phosphorylated SHIP2 on Y1135 localizes at focal adhesions and at the mitotic spindle in cancer cell lines. *Cell Signal* 2014; **26**: 1193–1203.

64 Kassenbrock CK, Anderson SM. Regulation of Ubiquitin Protein Ligase Activity in c-Cbl by Phosphorylation-induced Conformational Change and Constitutive Activation by Tyrosine to Glutamate Point Mutations. *J Biol Chem* 2004; **279**: 28017–28027.

65 Sahay S, Pannucci NL, Mahon GM, Rodriguez PL, Megjugorac NJ, Kostenko EV *et al.* The RhoGEF domain of p210 Bcr-Abl activates RhoA and is required for transformation. *Oncogene* 2007; **27**: 2064–2071.

66 Frank DA. STAT signaling in the pathogenesis and treatment of cancer. *Mol Med* 1999; **5**: 432–456.

67 Mikita T, Campbell D, Wu P, Williamson K, Schindler U. Requirements for interleukin-4-induced gene expression and functional characterization of Stat6. *Mol Cell Biol* 1996; **16**: 5811–5820.

68 Kornfeld J-W, Grebien F, Kerenyi MA, Friedbichler K, Kovacic B, Zankl B *et al.* The different functions of Stat5 and chromatin alteration through Stat5 proteins. *Front Biosci J Virtual Libr* 2008; **13**: 6237–6254.

69 Kee BL, Quong MW, Murre C. E2A proteins: essential regulators at multiple stages of B-cell development. *Immunol Rev* 2000; **175**: 138–149.

70 Ratliff ML, Mishra M, Frank MB, Guthridge JM, Webb CF. The Transcription Factor ARID3a Is Important for In Vitro Differentiation of Human Hematopoietic Progenitors. *J Immunol* 2016; **196**: 614–623.

71 Ichikawa M, Yoshimi A, Nakagawa M, Nishimoto N, Watanabe-okochi N, Kurokawa M. A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int J Hematol* 2013; **97**: 726–34.

72 Lam K, Zhang D-E. RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis. *Front Biosci J Virtual Libr* 2012; **17**: 1120–1139.

73 Kim DI, Cutler JA, Na CH, Reckel S, Renuse S, Madugundu AK *et al.* BioSITe: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation. *J Proteome Res* 2018; **17**: 759–769.

74 Chaiet L, Wolf FJ. The properties of streptavidin, a biotin-binding protein produced by Streptomycetes. *Arch Biochem Biophys* 1964; **106**: 1–5.

75 Cutler JA, Tahir R, Sreenivasamurthy SK, Mitchell C, Renuse S, Nirujogi RS *et al.* Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. *Leukemia* 2017. doi:10.1038/leu.2017.61.

76 Reckel S, Hamelin R, Georgeon S, Armand F, Jolliet Q, Chiappe D *et al.* Differential signaling networks of Bcr–Abl p210 and p190 kinases in leukemia cells defined by functional proteomics. *Leukemia* 2017. doi:10.1038/leu.2017.36.

77 Weisswange I, Newsome TP, Schleich S, Way M. The rate of N-WASP exchange limits the extent of ARP2/3-complex-dependent actin-based motility. *Nature* 2009; **458**: 87–91.

78 Tomas A, Futter CE, Eden ER. EGF receptor trafficking: consequences for signaling and cancer. *Trends Cell Biol* 2014; **24**: 26–34.

79 Dunn GP, Cheung HW, Agarwalla PK, Thomas S, Zektser Y, Karst AM *et al.* In vivo multiplexed interrogation of amplified genes identifies GAB2 as an ovarian cancer oncogene. *Proc Natl Acad Sci U S A* 2014; **111**: 1102–1107.

80 Zheng Y, Zhang C, Croucher DR, Soliman MA, St-Denis N, Pasculescu A *et al.* Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature* 2013; **499**: 166–171.

81 Brehme M, Hantschel O, Colinge J, Kaupe I, Planyavsky M, Köcher T *et al.* Charting the molecular network of the drug target Bcr-Abl. *Proc Natl Acad Sci* 2009; **106**: 7414–7419.

82 Schaller-Schönitz M, Barzan D, Williamson AJK, Griffiths JR, Dallmann I, Battmer K *et al.* BCR-ABL affects STAT5A and STAT5B differentially. *PloS One* 2014; **9**: e97243.

83 Hantschel O, Warsch W, Eckelhart E, Kaupe I, Grebien F, Wagner K-U *et al.* BCR-ABL uncouples canonical JAK2-STAT5 signaling in chronic myeloid leukemia. *Nat Chem Biol* 2012; **8**: 285–293.

84 Lee S-Y, Kang M-G, Shin S, Kwak C, Kwon T, Seo JK *et al.* Architecture Mapping of the Inner Mitochondrial Membrane Proteome by Chemical Tools in Live Cells. *J Am Chem Soc* 2017; **139**: 3651–3662.

85 Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res* 2016; **44**: D1251-1257.

86 Pawlak KJ, Prasad M, Thomas JL, Whittal RM, Bose HS. Inner Mitochondrial Translocase Tim50 Interacts with 3β-Hydroxysteroid Dehydrogenase Type 2 to Regulate Adrenal and Gonadal Steroidogenesis. *J Biol Chem* 2011; **286**: 39130–39140.

87 de la Cruz L, Bajaj R, Becker S, Zweckstetter M. The intermembrane space domain of Tim23 is intrinsically disordered with a distinct binding region for presequences. *Protein Sci Publ Protein Soc* 2010; **19**: 2045–2054.

88 Wu M, Gu J, Guo R, Huang Y, Yang M. Structure of Mammalian Respiratory Supercomplex I1III2IV1. *Cell* 2016; **167**: 1598-1609.e10.

89 Floyd BJ, Wilkerson EM, Veling MT, Minogue CE, Xia C, Beebe ET *et al.* Mitochondrial Protein Interaction Mapping Identifies Regulators of Respiratory Chain Function. *Mol Cell* 2016; **63**: 621–632.

90 Hart GW, Housley MP, Slawson C. Cycling of O-linked β-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* 2007; **446**. doi:10.1038/nature05815.

91 Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. Cross Talk Between O-GlcNAcylation and Phosphorylation: Roles in Signaling, Transcription, and Chronic Disease. *Annu Rev Biochem* 2011; **80**: 825–858.

92 Love DC, Hanover JA. The Hexosamine Signaling Pathway: Deciphering the 'O-GlcNAc Code'. *Sci STKE* 2005; **2005**: re13–re13.

93 Ning X, Guo J, Wolfert MA, Boons G-J. Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. *Angew Chem Int Ed* 2008; **47**: 2253–2255.

94 Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew Chem Int Ed* 2002; **41**: 2596–2599.

95 Wang Z, Udeshi ND, Slawson C, Compton PD, Sakabe K, Cheung WD *et al.* Extensive Crosstalk Between O-GlcNAcylation and Phosphorylation Regulates Cytokinesis. *Sci Signal* 2010; **3**: ra2–ra2.

96 Myers SA, Daou S, Affar EB, Burlingame A. Electron transfer dissociation (ETD): The mass spectrometric breakthrough essential for O-GlcNAc protein site assignments—a study of the O-GlcNAcylated protein Host Cell Factor C1. *PROTEOMICS* 2013; **13**: 982–991.

97 Mackay DR, Elgort SW, Ullman KS. The Nucleoporin Nup153 Has Separable Roles in Both Early Mitotic Progression and the Resolution of Mitosis. *Mol Biol Cell* 2009; **20**: 1652–1660.

98 Niessen S, Dix MM, Barbas S, Potter ZE, Lu S, Brodsky O *et al.* Proteome-wide Map of Targets of T790M-EGFR-Directed Covalent Inhibitors. *Cell Chem Biol* 2017; **0**. doi:10.1016/j.chembiol.2017.08.017.

99 Bar-Peled L, Kemper EK, Suciu RM, Vinogradova EV, Backus KM, Horning BD *et al.* Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. *Cell* 2017; **171**: 696-709.e23.

100 Strobl B, Moriggl R. Editorial: Recovery from chemotherapy depends on STAT1 for replenishment of B lymphopoiesis. *J Leukoc Biol* 2014; **95**: 849–851.

101 Hoelbl A, Schuster C, Kovacic B, Zhu B, Wickre M, Hoelzl MA *et al.* Stat5 is indispensable for the maintenance of bcr/abl-positive leukaemia. *EMBO Mol Med* 2010; **2**: 98–110.

102 Warsch W, Kollmann K, Eckelhart E, Fajmann S, Cerny-Reiterer S, Hölbl A *et al.* High STAT5 levels mediate imatinib resistance and indicate disease progression in chronic myeloid leukemia. *Blood* 2011; **117**: 3409–3420.

103 Daubon T, Chasseriau J, El Ali A, Rivet J, Kitzis A, Constantin B *et al.* Differential motility of p190bcr-abl- and p210bcr-abl-expressing cells: respective roles of Vav and Bcr-Abl GEFs. *Oncogene* 2008; **27**: 2673–2685.

104 Rochelle T, Daubon T, Troys MV, Harnois T, Waterschoot D, Ampe C *et al.* p210bcr-abl induces amoeboid motility by recruiting ADF/destrin through RhoA/ROCK1. *FASEB J* 2013; **27**: 123–134.

105 Thien CBF, Langdon WY. Cbl: many adaptations to regulate protein tyrosine kinases. *Nat Rev Mol Cell Biol* 2001; **2**: 294–307.

106 Rohrschneider LR, Fuller JF, Wolf I, Liu Y, Lucas DM. Structure, function, and biology of SHIP proteins. *Genes Dev* 2000; **14**: 505–520.

107 Chen X, Ren L, Kim S, Carpino N, Daniel JL, Kunapuli SP *et al.* Determination of the Substrate Specificity of Protein-tyrosine Phosphatase TULA-2 and Identification of Syk as a TULA-2 Substrate. *J Biol Chem* 2010; **285**: 31268–31276.

108 Hu Y, Liu Y, Pelletier S, Buchdunger E, Warmuth M, Fabbro D *et al.* Requirement of Src kinases Lyn, Hck and Fgr for BCR-ABL1-induced B-lymphoblastic leukemia but not chronic myeloid leukemia. *Nat Genet* 2004; **36**: 453–461.

109 Kovacic B, Hoelbl A, Litos G, Alacakaptan M, Schuster C, Fischhuber KM *et al.* Diverging fates of cells of origin in acute and chronic leukaemia: Cells of origin of BCR/ABL $^{+}$ CML and B-ALL. *EMBO Mol Med* 2012; **4**: 283–297.

110 Reckel S, Hamelin R, Georgeon S, Armand F, Jolliet Q, Chiappe D *et al.* Differential signaling networks of Bcr-Abl p210 and p190 kinases in leukemia cells defined by functional proteomics. *Leukemia* 2017; **31**: 1502–1512.

CURRICULUM VITAE

JEVON ANDREW CUTLER

529 S Port Street, jevon@jhmi.edu
Baltimore, Maryland, 21224 (503) 915-4557

EDUCATION

Ph.D.	2011-*present*	Predoctoral Human Genetics Training Program, Department of Biological Chemistry, Johns Hopkins University School of Medicine (JHU SOM), Baltimore, MD U.S.A Mentor: Karen L. Reddy, Ph.D., and Akhilesh Pandey, M.D., Ph.D Thesis Focus: Signaling in leukemia and improvement on proximity dependent labeling technologies. *Anticipated graduation May 2018*
B.A.	1998-2003	Political Science (Distinction) University of Oregon, Eugene, OR U.S.A.
	2006-2010	Post Baccalaureate Chemistry Portland State University, Portland, OR U.S.A.

RESEARCH EXPERIENCE

Research Rotation	2011	Laboratory of Dr. Steven Leach, Ph.D JHU SOM Department of Surgical Oncology
Staff analyst	2003-2011	*HematoLogics Inc.*, Seattle, WA Project Title: "Improving myelodysplastic syndrome diagnostics by integrating flow cytometry and cytogenetics."

PUBLICATIONS

Kim DI,* **Cutler JA,*** Na CH, Reckel S, Renuse S, Madugundu A, Tahir R, Goldschmidt HL, Reddy K, Huganir R, Wu X, Zachara NE, Hantschel O, Pandey A, (2017). BioSITe: A method for direct detection of site-specific biotinylation. *Journal of Proteome Research.* PMID: 29249144

Cutler JA, Tahir R, Huang TC, Patil AH, Reddy K, Pandey A (2017). Proteomic analysis of relapse in infant MLL-r leukemia reveals lineage infidelity and patient heterogeneity. *Manuscript in preparation.*

Cutler JA, Tahir R, Sreenivasamurthy SK, Mitchell C, Renuse S, Nirujogi RS, Patil AH, Heydarian M, Wong X, Wu X, Huang TC, Kim MS, Reddy K, Pandey A, (2017). Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. *Leukemia.* PMID: 28210003.

Huang TC, **Cutler JA**, Bharne S, Zhong J, Weinstock D, Tyner J, Gojo I, Civin C, Pandey A, (2015). Integrated analysis of CRLF2 signaling in acute lymphoblastic leukemia identifies Polo-like kinase 1 as a potential therapeutic target. *Leukemia Lymphoma*. PMID: 25213184

Mitchell C, Getnet D, Kim MS, Srikanth S, Kumar P, Pinto Iwasaki M, Huang T, Shaw P, Wu X, Zhong J, Chaerkady R, Muthusamy B, Nirujogi RS, Sahasrabuddhe N, Raju R, Bowman C, Danilova L, **Cutler JA**, Kelkar D, Drake C, Prasad TS, Marchionni L, Marukami P, Scott A, Shi L, Thierry-Mieg J, Thierry-Mieg D, Irizarry R, Wang C, Cope L, Ishihama Y, Gowda HC, and Pandey A, (2015). A multi-omic analysis of human naïve CD4+ T cells. *BMC Systems Biology*. PMID: 26542228

Heydarian M, Luperchio TR, **Cutler JA**, Mitchell CJ, Kim MS, Pandey A, Sollner-Webb B, Reddy K, (2014). Prediction of Gene Activity in Early B Cell Development Based on an Integrative Multi-Omics Analysis (2014). *J. Proteomics and Bioinformatics.* PMCID: PMC4276347.

Loosdrecht AA van de, Kern W, Porta MGD, Alhan C, Balleisen JS, Bettelheim P, Bowen DT, Burbury K, Eidenschink L, Cazzola M, Chu SSC, Cullen M, **Cutler JA**, Dräger AM, Feuillard J, Fenaux P, Font P, Germing U, Haase D, Hellström-Lindberg E, Johansson U, Kordasti S, Loken MR, Malcovati L, Marvelde JG te, Matarraz S, Milne T, Moshaver B, Mufti GJ, Nikolova V, Ogata K, Oelschlaegel U, Orfao A, Ossenkoppele GJ, Porwit A, Platzbecker U, Preijers F, Psarra K, Richards SJ, Subirá D, Seymour JF, Tindell V, Vallespi T, Valent P, Velden VHJ van der, Wells DA, Witte TM de, Zettl F, Béné, MC, Westers TM, (2013). Rationale for the clinical application of flow cytometry in patients with myelodysplastic syndromes: position paper of an International Consortium and the European LeukemiaNet Working Group. *Leukemia Lymphoma* PMID: 22916713

Westers TM, Ireland R, Kern W, Alhan C, Balleisen JS, Bettelheim P, Burbury K, Cullen M, **Cutler JA,** Della Porta MG, Dräger AM, Feuillard J, Font P, Germing U, Haase D, Johansson U, Kordasti S, Loken MR, Malcovati L, te Marvelde J.G, Matarraz S, Milne T, Moshaver B, Mufti GJ, Ogata K, Orfao A, Porwit A, Psarra K, Richards SJ, Subirá D, Tindell V, Vallespi T, Valent P, van der Velden VHJ, de Witte TM, Wells DA, Zettl F, Béné MC, van de Loosdrecht AA, (2012). Standardization of flow cytometry in myelodysplastic syndromes: a report from an international consortium and the European Leukemia Net Working Group. *Leukemia* PMID: 22307178

Cutler JA, Wells DA, van de Loosdrecht AA, de Baca ME, Kalnoski MH, Zehentner BK, Eidenschink L, Ghirardelli KM, Biggerstaff JS, Loken MR, (2011). Phenotypic abnormalities strongly reflect genotype in patients with unexplained cytopenias. *Cytometry B Clinical Cytometry*. PMID: 21520402

Zehentner BK, **Cutler JA**, Fritschle WK, Bennington RK, Wentzel C, Smading SR, Jeffery EW, Wells DA, Loken MR, (2011). A minority of concurrent monoclonal lymphocytes and plasmacytic cells sharing light chains are genetically related in putative lymphoplasmacytic lymphoma. *Leukemia Research*. PMID: 21752465
***Co-first author**

MENTORSHIP/TEACHING EXPERIENCE

2013 Graduate Teaching Assistant
Advanced Topics in Human Genetics (Fall), JHUSOM
Assisted in course organization and homework assignment grading in this core class for the Human Genetics program.

2004-2005 Workshop Leader
Organic Chemistry, Portland State University, Portland, OR
Organized weekly workshops for students taking organic chemistry at Portland State University by teaching material to reinforce lectures and helping students to perform better on exams.

PRESENTATIONS

Cutler, JA, Tahir, R, Kostadinov, R, Sekhar, R, Huang, TC, Brown, P, Heydarian, M, Pandey, A, Reddy, K.(2016) Proteomic/Transcriptomic Signatures of Infant MLL-r Rearranged B-ALL at Diagnosis and Relapse Reveal Lineage Plasticity and Diagnostic Heterogeneity **Poster Presentation.** December 1-4, 2016 ASH Annual Meeting & exposition. 128 (22), 2697-2697

Cutler, JA, Tahir, R, Han, J, Nirujogi, RS, Huang, TC, Wong, X, Mallampati, S, Sun, X, Brown, P, Reddy, K, Pandey, A. (2015) Differential Signaling through p190 and p210 Forms of BCR-ABL Fusion Proteins Revealed By Proteomic Analysis **Poster Presentation.** December 5-8, 2015 ASH Annual Meeting & exposition. Blood 126 (23), 3651-3651

Title:	BioSITe: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation
Author:	Dae In Kim, Jevon A. Cutler, Chan Hyun Na, et al
Publication:	Journal of Proteome Research
Publisher:	American Chemical Society
Date:	Feb 1, 2018

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK CLOSE WINDOW

SPRINGER NATURE

Title:	Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis
Author:	J A Cutler, R Tahir, S K Sreenivasamurthy, C Mitchell, S Renuse et al.
Publication:	Leukemia
Publisher:	Springer Nature
Date:	Feb 17, 2017

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select no to the question "Are you the Author of this Springer Nature content?".
Ownership of copyright in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:
To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).
The author and any academic institution, where they work, at the time may reproduce the contribution for the purpose of course teaching.
To reuse figures or tables created by the Author and contained in the Contribution in oral presentations and other works created by them.
To post a copy of the contribution as accepted for publication after peer review (in locked Word processing file, of a PDF version thereof) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the contribution on the publisher's website.
Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read Springer Nature's online author reuse guidelines.

For full paper portion: Authors of original research papers published by Springer Nature are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v1.0

BACK CLOSE WINDOW